

November 2018

Testing the Validity and Utility of a Career and Technical Education Programs of Study Evaluation Instrument

Alisha D. Hyslop
University of South Florida, ahyslop@actonline.org

Follow this and additional works at: <https://digitalcommons.usf.edu/etd>



Part of the [Other Education Commons](#)

Scholar Commons Citation

Hyslop, Alisha D., "Testing the Validity and Utility of a Career and Technical Education Programs of Study Evaluation Instrument" (2018). *USF Tampa Graduate Theses and Dissertations*.
<https://digitalcommons.usf.edu/etd/7524>

This Dissertation is brought to you for free and open access by the USF Graduate Theses and Dissertations at Digital Commons @ University of South Florida. It has been accepted for inclusion in USF Tampa Graduate Theses and Dissertations by an authorized administrator of Digital Commons @ University of South Florida. For more information, please contact digitalcommons@usf.edu.

Testing the Validity and Utility of a Career and Technical Education Programs of Study
Evaluation Instrument

by

Alisha D. Hyslop

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
in Curriculum and Instruction
with a concentration in Career and Workforce Education
Department of Leadership, Counseling, Adult, Career and Higher Education
College of Education
University of South Florida

Major Professor: Victor Hernandez, Ph.D.
Co-Major Professor: Liliana Rodriguez, Ph.D.
Edward Fletcher, Ph.D.
Robert Dedrick, Ph.D.

Date of Approval:
October 25, 2018

Keywords: CTE, quality, Association for Career and Technical Education, program theory

Copyright © 2018, Alisha D. Hyslop

DEDICATION

This dissertation is dedicated to my husband, Jason, for his unwavering support as I undertook this journey; to my children, Mackenzie and Cooper, who I hope have learned a bit about the dedication necessary to accomplish one's goals; and to the hundreds of thousands of CTE educators around the country who deliver high-quality programs to meet student needs every day!

ACKNOWLEDGMENTS

I would like to express my sincere thanks and appreciation to my doctoral committee. First, to my major professor Dr. Victor Hernandez, who patiently guided me through this journey from day one, when I sat in his office and expressed interest in joining a cohort just a few days before the deadline. I am extremely grateful for his calm demeanor and thoughtful advice over the many years of my work. I would also like to thank my co-major professor, Dr. Liliana Rodriguez, who introduced me to the formal practice of evaluation and shaped the entire focus of this work. She was always ready with a kind word of encouragement and her passion for students was evident in every interaction. Much appreciation also goes to Dr. Edward Fletcher, whose questions and advice were extremely helpful throughout my journey. And last, but certainly not least, the statistical expertise of Dr. Robert Dedrick was critical to the success of this project—especially as he jumped in to help with additional data analysis to answer important questions.

In addition to my formal committee, I would like to acknowledge the many CTE researchers around the country who provided assistance during my doctoral journey, particularly Dr. James Stone and Dr. Pradeep Kotamraju. The advice they offered throughout the process was invaluable, and I greatly appreciated their time, expertise, and willingness to share! Finally, I would like to thank my research partner on the first phase of ACTE's high-quality CTE research, which led to this project—Catherine Imperatore. This work would not have been possible without her focus and expertise, and continued advice throughout the project!

TABLE OF CONTENTS

LIST OF TABLES	iv
LIST OF FIGURES	v
ABSTRACT	vi
CHAPTER ONE: INTRODUCTION.....	1
Problem Statement	1
Purpose of the Study	8
Research Questions	8
Conceptual Framework	10
Significance of the Study	12
Limitations	12
Definitions of Key Terms	13
CHAPTER 2: LITERATURE REVIEW	17
Career and Technical Education	17
History of CTE.....	17
Current Legislation	20
New Constructs.....	23
Career Clusters.....	23
Programs of Study.....	23
Career Pathways.....	25
Overview of Evaluation	26
Formative vs. summative evaluation.	26
Evaluation approaches.	27
Evaluation models.....	28
Evaluation of CTE	29
Best practices identification.....	30
Perkins accountability.....	32
National Assessment of Career and Technical Education	33
State efforts	34
National frameworks.....	35
ACTE High-quality CTE Program of Study Research	39
Analysis of current quality frameworks.....	40
Development of ACTE Framework.....	43
Conceptual Framework.....	45
Program theory-driven evaluation.	45
Validity of evaluation instruments.....	47

Translation validity	48
Criterion-related validity.....	49
Summary of Literature Review.....	50
CHAPTER 3: METHODOLOGY	51
Research Design.....	51
Phase One.....	53
Program sample	53
Data	55
Instrumentation	56
Data collection procedures.....	58
Data analysis	58
Question one	59
Question two	60
Phase Two.....	60
Sample.....	60
Instrumentation	61
Data collection procedures.....	61
Data analysis	62
CHAPTER 4: RESULTS	64
Phase One.....	64
Description of sample.	64
Evaluation instrument results.....	67
Validity of evaluation instrument.	68
Utility of the evaluation instrument	72
Ease of use	72
Usefulness of instrument.....	73
Implications for future use	75
Practicality across CTE program models.....	76
Phase Two.....	77
Description of sample	78
Emergent themes.....	78
The instrument is seen as a valid measure of program quality	79
Administrators and educators may have different perspectives	80
This instrument fills a gap in available quality metrics	82
The use of this instrument lends credibility to participating programs	83
The instrument is an effective program improvement tool across the CTE spectrum	84
CHAPTER 5: DISCUSSION.....	87
Study Overview	87
Summary of Results.....	88
Evaluation instrument results.....	88
Research question one.....	88
Research question two	90

Discussion and Conclusions	92
Results on the instrument.....	93
Validity of evaluation instrument	93
Utility of evaluation instrument	94
Limitations	96
Next Steps in ACTE’s High-quality CTE Initiative	98
Implications for Policy.....	100
Implications for Future Research.....	101
REFERENCES	104
APPENDICES	110
Appendix A: Institutional Review Board (IRB) Letter.....	111
Appendix B: Conflict of Interest Disclosure Approval and Management Plan.....	112
Appendix C: Evaluation Instrument Rating Scale	116
Appendix D: Evaluation Instrument	117
Appendix E: Feedback Questionnaire	138
Appendix F: Phase Two Interview Protocol	144
Appendix G: Results From Individual Element Correlations.....	145

LIST OF TABLES

Table 2.1: Criteria from Exemplary Career and Technical Education Programs: Secondary-Level User’s Guide to Program Improvement.....	31
Table 2.2: 2006 Carl D. Perkins Act Accountability Indicators	32
Table 2.3: Program of Study Framework Components and Subcomponents	36
Table 2.4: Criteria for Linked Learning Pathway Quality Review and Continuous Improvement	38
Table 4.1: Program Evaluation Form Participants by Career Cluster	66
Table 4.2: Program Evaluation Instrument Results: Information on Scores for Criteria Within Each Element	69
Table 4.3: Correlation Results	70
Table 4.4: Evaluation Instrument Ease of Use Ratings	73
Table 4.5: Insights into Improving Programs	74

LIST OF FIGURES

Figure 2.1: Linear program theory model.....	46
Figure 3.1: Two-phase, sequential mixed methods research design.....	52
Figure 4.1: Suggestions for Future Use	76

ABSTRACT

The purpose of this study was to build on the research already completed by the Association for Career and Technical Education (ACTE) in the development of ACTE's quality Program of Study Framework and accompanying evaluation instrument to examine the validity of the instrument's results and its utility when used independently by local educators in a pilot test. The study used a two-phase mixed-methods design. In the first phase, 39 participating programs were evaluated using the instrument, consisting of 102 items organized into 12 elements. Representatives of participating programs, either teachers or administrators, then provided quantitative and qualitative feedback about the instrument, and 23 of the programs submitted Perkins program performance data. Results on the program evaluation instrument were correlated with the Perkins performance data to determine the criterion validity of the instrument. Correlations were significant for two of the performance measures, student technical skill attainment and completion, and positive but not statistically significant for student placement. Results related to the utility of the instrument, including for ease of use, practicality across programs, and usefulness of information, were all positive, and participants provided suggestions to further increase the possibility of wide scale future use. In Phase Two of the study, results of Phase One were used to inform in-depth interviews with a sample of five participants. These interviews provided further insight on the validity and utility questions guiding the study. The study is significant because it helps to more clearly define the characteristics of a high-quality CTE program and provide a standardized way to determine and improve the quality of CTE programs for millions of students around the country.

CHAPTER ONE: INTRODUCTION

Recent federal policies, coupled with state and national economic concerns, have put pressure on all publicly funded programs to produce results (Hollenbeck, 2011). Policymakers at the local, state, and federal levels all want to know that their investments are worthwhile and that practitioners are using scarce resources in the most efficient manner. To help produce these desired results, there has been a renewed focus within public education on identifying specific practices that lead to positive outcomes and spreading such practices across the country (U.S. Department of Education, 2013). Essentially, policymakers can no longer be satisfied with simply knowing whether education funding streams “work,” they also must understand the key elements of how and why programs work so that those specific elements can be replicated more widely and achieve desired results (Hess & Little, 2015).

Greater attention to the identification and scalability of best practices can be seen throughout education, but is particularly necessary within the field of career and technical education (CTE) (Duncan, 2011). Unfortunately, CTE lacks a set of clearly defined best practices and a common definition of “what works” to improve student achievement, making the task difficult. This study builds on previous work by the Association for Career and Technical Education (ACTE) to address the deficiency of program-quality criteria by testing the validity and utility of ACTE’s new program evaluation tool.

Problem Statement

Despite large enrollments and endorsements from many business and industry leaders, high program costs and the recent increased focus on core academics have meant CTE programs,

even more than other publicly funded education activities, are facing growing demands from policymakers to improve outcomes and provide verification of student achievement. Without high-quality program evaluation metrics, it is becoming increasingly difficult for programs to justify their existence (Duncan, 2011).

CTE programs that prepare students for their post-education future have been a part of public education in the United States since the late 1800s, and have been supported by direct federal resources since the Smith Hughes Act was passed in 1917 (Barlow, 1976). Over the years, these programs have evolved from preparing students solely for entry-level jobs to imparting a rigorous mix of academic, technical, and employability skills that allow students to pursue further education and a wide variety of careers (ACTE, n.d.). Now, 94% of the country's high school students enroll in at least one CTE course, with approximately 20% participating in a sequence of three or more courses (U.S. Department of Education, National Center for Education Statistics, 2009). Approximately 12 million postsecondary students participate as well (ACTE, n.d.).

Over the years, new program models and delivery structures have emerged to support the broader mission of CTE. In the late 1990s, the Career Clusters model helped to organize career possibilities into 16 broad categories, each containing more specific pathways. The foundational idea behind the Clusters movement was to better organize content and standards so that students learned academic and technical skills along an integrated continuum. The new Career Cluster framework allowed for more career exploration activities and the ability to expose students to a broader array of career possibilities within their area of interest, rather than a narrow, single-occupation approach to instruction (Ruffing, 2006).

Further, to better connect secondary and postsecondary education, the concept of a program of study was introduced in the Carl D. Perkins Career and Technical Education Act of 2006. A program of study is a sequence of academic and technical courses that spans both secondary and postsecondary education and leads to a credential. Specifically, the law referenced “coherent and rigorous content aligned with challenging academic standards and relevant career and technical content in a coordinated, non-duplicative progression of courses that align secondary education with postsecondary education to adequately prepare students to succeed in postsecondary education” (Section 122(c)(1)(A)(ii)). In practice, programs of study have driven the CTE community to provide clearer, more cohesive program offerings to students, to better connect secondary and postsecondary educational levels, and to more carefully consider in-demand credentials.

However, the breadth of the CTE enterprise and its long, evolving history are not enough to sustain it into the future. One specific criticism has been the failure to share and grow evidence-based practices that would help all programs produce more desired outcomes (U.S. Department of Education, Office of Vocational and Adult Education, 2012). This criticism relates both to the need to better evaluate CTE programs, as well as to share these results in a way that will spread best practices.

Foundations of the first step in the work to better evaluate CTE programs—identifying programs with positive outcomes—are already in place. The 1998 Carl D. Perkins Vocational and Technical Education Act first implemented specific federal performance indicators for CTE programs, and Congress strengthened these requirements during the law’s 2006 reauthorization (ACTE, 2006). The 2006 Carl D. Perkins Career and Technical Education Act requires states and local school districts to report performance information on eight separate performance

measures for secondary programs, and six measures at the postsecondary level (U.S. Department of Education, Office of Career, Technical, and Adult Education, Division of Academic and Technical Education, n.d.a). A slightly revised set of performance indicators was included in the Perkins Act's 2018 reauthorization, which will go into effect on July 1, 2019, but the themes are continued (Advance CTE & ACTE, 2018).

However, it is clear from the criticisms that the field of CTE must do more to evaluate programs in a way that not only identifies positive student outcomes, but also allows the practices that led to these outcomes to be more easily recognized and such information to be more easily shared across the country. While the Perkins performance measures can form a basis for identifying high- and low-performing programs, they are insufficient for answering underlying questions about how or why that level of achievement was attained. On their own, they do not provide the information necessary to identify and replicate positive practices or to allow program leaders to determine appropriate technical assistance for targeted program improvement. For these reasons, more extensive evaluation instruments to examine program elements and measure progress are necessary. These instruments must serve to help unpack Suchman's (1967) "black box" to determine how programs produce results.

Over the past few years, some states and national organizations have begun the process of looking at program inputs, activities, and outputs in an effort to determine which programs are on track for success. However, the depth, contents, quality, and even purpose of their evaluation instruments vary tremendously, and their use is inconsistent. For example, some states have implemented evaluation tools merely to monitor program compliance with federal and state laws—which, while a valuable purpose, does not meet the broader goal of ensuring program quality and scalability of best practices. Others have developed their evaluation instruments to

provide information more useful for program improvement, but there is no common understanding of which program elements are most important or even necessary to ensure positive student outcomes (Imperatore & Hyslop, 2015a).

Examples of evaluation tools at the national level generally fall into three categories: quality frameworks that evaluate specific types of CTE programs, those that relate to entire institutions, and those that apply to broader career pathways systems (Imperatore & Hyslop, 2015a). In the first category are efforts like the National Career Academy Coalition's quality framework, which is only applicable to programs delivered specifically through the career academy model (National Career Academy Coalition, 2013), and the Department of Education's Program of Study Design Framework, which contains 10 components and related subcomponents that the federal Office of Career, Technical, and Adult Education (formally OVAE) suggests "support the development and implementation of effective programs of study" within CTE (OVAE, 2010, p. 1). Unfortunately, the Program of Study Design Framework includes elements that vary in level of detail and audience, limiting its utility for local practitioners (Imperatore & Hyslop, 2015a).

The second category of national frameworks is those evaluating entire institutions, including those from accrediting bodies, as well as national organizations such as High Schools that Work. Finally, at the broader systems level, there have been a number of new efforts to incorporate CTE programs into the growing career pathways movement, and a number of quality-related frameworks related to these career pathway systems. These range from government documents, such as the Department of Labor's six key elements of career pathways and Department of Health and Human Services' descriptive and service strategies, to the Criteria

and Indicators for a Quality State Career Pathways System from the Center for Law and Social Policy (Imperatore & Hyslop, 2015a).

The growing proliferation of CTE-related standards and frameworks has left the field without a clear definition of program quality, and opens the door for criticisms like those referenced earlier from federal policymakers. To begin to address this concern, the Association for Career and Technical Education (ACTE) launched its High Quality CTE Initiative in July of 2015 to “help synthesize the myriad voices that are part of the dialogue on high-quality CTE” (Imperatore & Hyslop, 2015a, p. 1). This multi-phase project is designed “to identify a comprehensive, research-based quality CTE program of study framework, test the framework and integrate it into [the Association’s] efforts to recognize and disseminate information on best practices within CTE” (Imperatore & Hyslop, 2015a, p. 1).

Four pieces have already been publically released from ACTE under this initiative: *Defining High-quality CTE: Contemporary Perspectives on CTE Quality*, in July 2015; *Defining High-quality CTE: Quality CTE Program of Study Framework, Draft Version 1.0*, in November 2015; *Defining High-quality CTE: Quality CTE Program of Study Framework, Draft Version 3.0*, in November 2016 (draft 2.0 of the referenced Framework was used for research purposes and not publically released); and *Defining High-quality CTE: Quality CTE Program of Study Framework, Beta Version 4.0. Version 4.0*, which will be referenced throughout this study as the “ACTE Framework” or the “Framework”

In the first paper, ACTE researchers catalogued and sorted the various program quality frameworks currently being used in the field, including those mentioned earlier, and then coded them based on key structural and content elements. This work led to the conclusion that there is a lack of clarity and consistency in the current definition of high-quality CTE, particularly as it

relates to individual local programs (Imperatore & Hyslop, 2015a). There was not a clear set of quality criteria that could be applied universally to specific CTE programs by local practitioners, either for program improvement or evaluation purposes.

From this initial work, ACTE researchers then set out to fill the identified gap by creating a comprehensive, research-based framework for local programs of study. The “program of study” was selected as the unit of analysis on which to base the framework due to its prominence in federal legislation and state and local implementation efforts related to CTE. It is also broad enough to encompass the diverse CTE delivery system, including programs in comprehensive high schools, magnet schools, area CTE centers, early college high schools, community and technical colleges, and more—while at the same time providing a foundational structure to the work (Personal Consultation, ACTE).

In November 2015, the *Draft Version 1.0* of the ACTE Framework was released. It contained 12 main elements, ranging from student assessment to career development. Under each element was a set of more specific criteria that further elaborate on the element’s implementation. ACTE then began the process of ensuring the content validity of the elements and criteria through a number of research strategies, including the use of additional research and literature reviews, and expert input provided through review, surveys and focus groups. These strategies were designed to build consensus around the elements of the Framework and the fact that these are indeed the elements necessary for a high-quality CTE program that will lead to desired student outcomes. The validation process included several rounds of revisions to incorporate input and standardize elements, and *Beta Version 4.0* of the Framework was released in the spring of 2017.

After this series of activities to establish content validity and the finalization of the beta version for testing, the Framework was adapted into a standard evaluation instrument that can be used to measure the quality of individual programs of study, not only for the purposes of program assessment and accountability, but also for continual program improvement. While all of the programmatic details are still being finalized, ACTE's goals include ensuring the evaluation instrument is a useful tool for program self-evaluation, as well as third-party assessment linked to the identification of best practices and recognition programs. The gap that remained in this effort was the need to pilot test the evaluation instrument to ensure that it is a valid and useful tool in a variety of CTE settings around the country.

Purpose of the Study

The purpose of this study was to build on the research already completed by ACTE in the development of the ACTE Framework and accompanying evaluation instrument to examine the validity of the instrument's results and its utility when used independently by local educators in a pilot test.

Evidence of the instrument's criterion validity—whether the included elements are appropriate measures of program quality, as related to already established program outcome measures—was examined. In addition, the study examined its utility, including ease of use, practicality, and usefulness of results to local practitioners. Within these questions, the resources, barriers and possibilities for scaling up the tool to widespread use were considered.

Research Questions

The specific research questions that were examined to achieve the purposes of the study are as follows:

1. Does the program evaluation instrument provide valid evidence of program quality?

2. Is the evaluation instrument a useful tool for local educators to assess the quality of their CTE programs?

For the purposes of these questions, “valid evidence of program quality” was measured primarily by correlations with established program outcome data based on the primary core indicators of performance as defined in the Carl D. Perkins Career and Technical Education Act of 2006, including students’ academic and technical skill attainment, completion, and placement rates as reported under that Act. These performance indicators are the required standard by which program quality is judged in the field and thus provided the most widely accepted set of program quality expectations.

Results on the evaluation instrument were correlated with program outcome data to provide evidence of its validity, focusing on the instrument’s predictive validity as a measure of criterion-related validity. Drost (2011) defines criterion-related validity as “the degree of correspondence between a test measure and one or more external referents (criteria), usually measured by their correlation” (p. 118). In this study, the “test measure” is the evaluation instrument in question, and the “one or more external referents” are the performance data. In addition, additional evidence of validity was collected to supplement the correlation data.

The measures of the instrument’s usefulness included its ease of use and practicality across education settings, and the usefulness of the results for both program improvement and evaluation purposes. Local educators included both faculty and administrators at the institution or district level who volunteered to pilot the evaluation tool on individual programs of study in their schools or institutions.

Conceptual Framework

This study was built upon the foundational elements of ACTE’s research to identify elements of a high-quality CTE program, outlined in Imperatore and Hyslop’s (2015a) *Defining High-quality CTE: Contemporary Perspectives on CTE Quality* and their framework and evaluation instrument development processes. Woven throughout this work, the constructs of program evaluation and instrument development then formed the basis of the conceptual framework for this study.

Evaluation can be defined as “the identification, clarification, and application of defensible criteria to determine an evaluation object’s value” (Fitzpatrick, Sanders, & Worthen, 2011, p. 7). ACTE’s early research on which this study is based set out to identify the “defensible criteria” that could be used to systematically identify and replicate quality within CTE programs of study. Specifically, principles of theory-driven evaluation guided the identification of these criteria by ACTE. Theory-driven evaluation is often aligned with the phenomenon of “unpacking the black box,” first proposed by Edward Suchman in 1967 to emphasize the importance of evaluating not only the overall effectiveness of social programs, but understanding the components of such programs.

Donaldson (2007) defines a three-step process for implementing theory-driven evaluation. First, an understanding of how a program is hypothesized to achieve its outcomes must be developed. Second, evaluation questions are then formulated, and then finally, answered. This process parallels the purpose of ACTE’s high-quality CTE initiative—first to validate the practices or “components” that will result in improved CTE student outcomes, and then to develop and implement an evaluation instrument based on such components.

Validity plays a key role in the first step of the theory-driven evaluation process and thus, the first step in the development of the ACTE Framework and evaluation instrument. During the initial phase of the ACTE Framework's development, researchers subjected it and the evaluation instrument to several tests of translation validity, including face validity and content validity. According to Drost (2011), "Translation validity attempts to assess the degree to which constructs are accurately translated into the operationalisation, using subjective judgment—face validity—and examining content domain—content validity" (p. 116). This was accomplished through rigorous content analysis of existing frameworks, literature reviews, and expert review and input.

Validity is also a key component of this phase of the project. The research questions of this study were based on the principles of criterion-related validity in particular, focused on establishing the predictive validity of the instrument. One of the major principles of a leading evaluation model by Kirkpatrick and Kirkpatrick (2009) is that when designing training programs, you must begin by thinking about the desired end results—level four, and work backward from that point to determine the behaviors that will lead to those results, the knowledge and skills that will improve those behaviors, and the environment that will best relay that information. This is an especially critical concept to apply to the evaluation of CTE programs, which must maintain a clear focus on desired program outcomes and determining the appropriate conditions to lead to those outcomes. In this vein, the criterion-related validity of the evaluation instrument was determined by correlating program input measures with desired end results, or program outcome measures.

Based on this conceptual framework focused on ACTE's early research, program-theory evaluation, and principles of validity, it was expected that the elements that were identified and

validated on their constructs in the initial stage of ACTE's research would lead to desired program outcomes measured by student performance. CTE programs that score higher on the evaluation instrument would have higher levels of student achievement, and programs that score lower overall on the evaluation instrument would have lower overall outcomes on student achievement measures.

Significance of the Study

This study makes several significant contributions to the field of CTE by helping to more clearly define the characteristics of a high-quality CTE program of study. Perhaps most important from a program survival standpoint, the development of a set of valid program-quality criteria and a single evaluation instrument provides the standardized evidence of program success that policymakers have demanded. This is a critical step for CTE programs that are under growing pressure to improve student outcomes, and could prove instrumental in requests for increased funding and other positive policy developments.

A single, valid instrument would also provide comparability to the state and local CTE program evaluation processes required by federal and state laws, and provide results useful for benchmarking and true program improvement. Exemplary programs could be more easily recognized for outstanding achievement in particular areas, and such information could be more easily shared in a manner designed to advance the performance of the entire system of CTE, ultimately increasing student access to high-quality programs and overall student achievement.

Limitations

In addition to the decisions made by the researcher that delimit the study, several limitations also exist. First, the diversity within the CTE field makes it very difficult for one instrument to accurately capture every single element that might be considered necessary for a

specific program to produce high-quality results. Some programs, particularly those dealing with specific, high-need populations such as students with disabilities or juvenile offenders, or with a more specialized subject, may not be able to use the evaluation instrument effectively without further refinement or additions. There are also many factors impacting student outcomes that will not be examined in this study, and could impact results if considered.

Definitions of Key Terms

Association for Career and Technical Education (ACTE). The Association for Career and Technical Education is a professional organization representing the full range of CTE professionals, from the middle school level to postsecondary education, including educators, administrators, career guidance counselors, business and industry partners, and related organizations. Members come from the local, state, and national level, and the organization works to provide professional development, leadership, and advocacy on behalf of the entire field.

Career Clusters®. Career Clusters are “an organizing tool for curriculum design and instruction (Advance CTE, 2016, para. 2)” within CTE. There are 16 national Career Clusters representing broad industries or occupational areas that are further divided into 79 career pathways. The 16 national Career Clusters include:

- Agriculture, Food & Natural Resources
- Architecture & Construction
- Arts, A/V Technology & Communications
- Business Management & Administration
- Education & Training
- Finance

- Government & Public Administration
- Health Science
- Hospitality & Tourism
- Human Services
- Information Technology
- Law, Public Safety, Corrections & Security
- Manufacturing
- Marketing
- Science, Technology, Engineering & Mathematics
- Transportation, Distribution & Logistics

Carl D. Perkins Act. The Carl D. Perkins Act is the primary federal legislation providing funding to CTE programs. The Act as currently being implemented was last reauthorized by Congress as the Carl D. Perkins Career and Technical Education Act of 2006, and prior to that was known as the Carl D. Perkins Vocational and Technical Education Act of 1998. Approximately \$1.2 billion is allocated through the law to states and then to local school districts and postsecondary institutions. Funds are used for local program improvement, including activities such as professional development, new equipment or curriculum, and CTE student support. A new piece of federal CTE legislation was passed in 2018, known as the Strengthening Career and Technical Education for the 21st Century Act, but will not go into effect until July 1, 2019, so is not yet relevant to this study (Advance CTE & ACTE, 2018).

Career pathways. For the purposes of this study, career pathways refer to a system of “connected education and training strategies and support services that enable individuals to secure industry relevant certification and obtain employment within an occupational area and to

advance to higher levels of future education and employment in that area” (U.S. Departments of Education, Labor, and Health and Human Services, 2012, pg. 1). Generally, they involve several levels of education and multiple agencies or private sector partners.

High-quality CTE. For the purposes of this study, ACTE’s 12 elements of high-quality CTE, as included in *Defining High-Quality CTE: Quality CTE Program of Study Framework, Beta Version 4.0* (Imperatore & Hyslop, 2017) are used to define program quality. These elements form the basis of the research project, and complete details can be found in Appendix D as part of the evaluation instrument. The 12 elements are:

1. Standards-aligned and Integrated Curriculum
2. Sequencing and Articulation
3. Student Assessment
4. Prepared and Effective Program Staff
5. Engaging Instruction
6. Access and Equity
7. Facilities and Equipment
8. Business and Community Partnerships
9. Career Development
10. Career Technical Student Organizations (CTSOs)
11. Work-based Learning
12. Data and Program Improvement

Program of study. The Carl D. Perkins Career and Technical Education Act of 2006 defines this term as options for students that:

- Incorporate secondary education and postsecondary education elements;

- Include coherent and rigorous content aligned with challenging academic standards and relevant career and technical content in a coordinated, non-duplicative progression of courses that align secondary education with postsecondary education adequately prepare students to succeed in postsecondary education;
- May include the opportunity for secondary education students to participate in dual or concurrent enrollment programs or other ways to acquire postsecondary education credits; and
- Lead to an industry-recognized credential or certificate at the postsecondary level, or an associate or baccalaureate degree (Section 122(c)(1)(A)).

CHAPTER 2: LITERATURE REVIEW

This study was built on efforts by ACTE to develop and validate a high-quality CTE framework and evaluation instrument. The study's specific purpose was to examine the evaluation instrument's validity and utility when used independently by local educators in a pilot test to measure CTE program quality. This chapter begins by presenting a conceptual overview of CTE, of general evaluation principles, and of specific efforts to evaluate the quality of CTE programs, including both formal and informal work at the national, state, and local levels. Next, early work by the Association for Career and Technical Education (ACTE) to identify current evaluation criteria, and gaps in that research, is summarized in detail as the foundation of the proposed study. Finally, the conceptual framework for the study situates it in theory-driven evaluation and program theory, as well as the principles of validity as they relate to evaluation and instrument development.

Career and Technical Education

History of CTE. CTE programs that prepare students for their post-education future have been a part of public education in the United States since the late 1800s, and have been supported by direct federal resources since the Smith-Hughes Act was passed in 1917 (Barlow, 1976). The Smith-Hughes Act of 1917 began the federal investment in secondary CTE, and marked the beginning of a more formalized system of what was then called “vocational education” in the United States. This legislation, coming on the heels of the industrial revolution and World War I, was designed to provide more skilled workers for industry by keeping students in high school and increasing job preparation (Gordon, 2014). It funded a limited range of

programs in agriculture, trade and industrial education, and home economics. Under the Smith-Hughes Act, vocational education was seen as separate from more academic studies, with requirements for separate state boards and limits on the amount of time students could engage in academic instruction (Hayward & Benson, 1993). Vocational education at this time was seen as an alternative to the more general, liberal arts curriculum that prevailed in most schools (Lynch, 2000).

Over the next five decades, vocational education programs, and the legislation that governed them, would grow and expand. Funding increased, and the types of fields supported did as well, growing to include programs in areas like practical nursing and marketing (Gordon, 2014). Beginning in 1963, the Vocational Education Act and its subsequent amendments expanded the purpose and focus of CTE even further, with a particular focus on those most in need of educational assistance. LaFollette (2011) described this expansion as for “persons of all ages in all communities.” Legislation—and programs—began to focus more on student needs and on accessibility to diverse and disadvantaged populations (Gordon, 2014). Set-asides for services for special population groups were hallmarks of the legislation in the 60s, 70s and 80s (Gordon, 2014), in many cases in response to the economic and social context of the time period (Hayward & Benson, 1993). The Carl D. Perkins Vocational Education Act of 1984 exemplified a growing dichotomy within CTE—two disparate goals of providing technical skills for the workforce and supporting students from special populations (Lynch, 2000; Hayward & Benson, 1993).

These disparate goals began to shift in the 1990s and 2000s, as the country’s education system was increasingly scrutinized. The release of the *A Nation at Risk: The Imperative for Educational Reform* report in the 80s (U.S. Department of Education, 1983) triggered a wave of

education reform efforts that eventually led to legislation emphasizing the integration of vocational education and academics, articulation across levels of education, closer partnerships between education with industry, and stronger accountability (Gordon, 2014). It was in this period that we began to see modern-day “CTE” emerge from the vocational education that existed for most of the century, and a decline in the isolation that had separated vocational and academic educational experiences.

From this shift came the career and technical education that is recognized today, imparting a rigorous mix of academic, technical, and employability skills that allow students to pursue further education and a wide variety of careers (ACTE, n.d.). Now, 94% of the country’s high school students enroll in at least one CTE course, with approximately 20% participating in a sequence of three or more courses (U.S. Department of Education, National Center for Education Statistics, 2009). Approximately 12 million postsecondary students participate as well (ACTE, n.d.). Courses are offered in middle schools, comprehensive high schools, magnet schools, career centers, community and technical colleges, four-year universities, and other alternative settings. CTE programs are defined by the federal Carl D. Perkins Career and Technical Education Act of 2006 as:

Organized educational activities that—(A) offer a sequence of courses that—(i) provides individuals with coherent and rigorous content aligned with challenging academic standards and relevant technical knowledge and skills needed to prepare for further education and careers in current or emerging professions; (ii) provides technical skill proficiency, an industry-recognized credential, a certificate, or an associate degree; and (iii) may include prerequisite courses (other than a remedial course) that meet the requirements of this subparagraph; and (B) include competency-based applied learning

that contributes to the academic knowledge, higher-order reasoning and problem-solving skills, work attitudes, general employability skills, technical skills, and occupation-specific skills, and knowledge of all aspects of an industry, including entrepreneurship, of an individual. (Section 3(5))

Current legislation. The current major federal legislation governing CTE was signed into law in 2006 as the Carl D. Perkins Career and Technical Education Act of 2006. Structurally, it remains similar to prior legislation, with a Basic State Grant providing CTE resources to states via formula, that in turn provide formula resources to local school districts and postsecondary institutions (ACTE, 2006). There are five major themes exemplified in this legislation: accountability, coordination within the CTE community, academic and technical integration, connections between secondary and postsecondary education, and links to business and industry (ACTE, 2006). Many of the themes were manifested in response to the then recent passage of the No Child Left Behind Act of 2002, which required an intense focus on student academic achievement and accountability (Gordon, 2014), and built on reforms began in the 1990s.

Conversations about accountability and enhancing the academic achievement of CTE students dominated the reauthorization conversations leading up to the passage of this law, and led to new requirements such as local performance targets, alignment with academic standards, and programs of study (ACTE, 2006). Programs of study, described in more detail in the following section, are designed to more intentionally connect secondary CTE to postsecondary CTE, and allow more students to transition seamlessly into postsecondary education. In addition, even the change in terminology from “vocational education” to “career and technical

education” emphasized the increased rigor and prominence of CTE as a path for all students (ACTE, 2006).

The 2006 Act also provided even more flexibility for states and local programs, with expanded “uses of funds” language, options for pooling resources, and the option to combine funding streams to make the best use of limited resources (ACTE, 2006). While at times this flexibility is critical, it has also led to some of the criticisms of CTE as being too variable (U.S. Department of Education, Office of Planning, Evaluation and Policy Development Policy and Program Studies Service, 2014). Even with this variability, however, the federal law has been recognized as critical to driving state and local efforts to improve and enhance CTE programs, and a mutually influential interchange has developed between federal and state policy (Klein et al., 2014).

With more than a decade since the last federal legislation, Congress began work on a new CTE law, and finally approved that bill in July 2018. Discussions began several years ago, but became more fully developed in late 2015 and 2016, and continued over the next few years. The Senate Health, Education, Labor, and Pensions Committee released a set of bipartisan principles guiding the reauthorization that represent much of policymakers’ early thinking:

1. Make it easier for States and locals to run their CTE programs to serve all students who desire to gain access to CTE coursework, including students with disabilities.
2. Increase access to, and support of, career counseling for all CTE students.
3. Maintain CTE as a formula program.
4. Align with ESEA and WIOA (where applicable) to improve the efficiency and effectiveness of the education and workforce development programs.

5. Support the expansion of public/private collaborations with secondary and postsecondary programs, including alignment with State or locally-determined in-demand industries and occupations.
6. Support efforts to integrate into and strengthen career pathways at the state and local levels.
7. Address unfunded programs.
8. Improve evaluation and research to support innovation and best practices. (Hyslop, 2015)

Reauthorization legislation was actually passed in the House in 2016, although work could not be completed in the Senate before Congress adjourned for the year. The Strengthening Career and Technical Education for the 21st Century Act, H.R. 5587 (2016), built on many of the Senate's themes, but also emphasized increased state and local control and the devolution of the federal role in CTE. Congress resumed reauthorization efforts in 2017, and the House again passed its bill, that year as H.R. 2353. The Senate picked up this bill during the summer of 2018 and passed an amended version, which then went back to the House for final passage before being signed into law. Since this new law, Public Law 115-224, does not go into effect until July 1, 2019, the current 2006 Perkins Act remains the standard for CTE programs nationwide (Advance CTE & ACTE, 2018).

The delay in federal reauthorization opened the door for more state activity in the interim. States have enacted an increasing number of CTE-related policies in recent years, building on the themes that have emerged in reauthorization conversations. In 2015 alone, there were more than 125 pieces of legislation, board rules, executive actions, and budget provisions enacted in states related to CTE (ACTE & Advance CTE, 2016). These state policy actions most commonly fell

into the categories of funding, governance, industry connections and work-based learning, career/academic guidance and awareness, and dual and concurrent enrollment.

New constructs. Over the years, new program models and delivery structures have emerged to support the broader mission of CTE.

Career Clusters. In the late 1990s, the Career Clusters model helped to organize career possibilities into 16 broad categories, each containing more specific pathways (now numbering 79 in total). The Clusters were developed out of work by the U.S. Department of Education and several other groups on skill standards, and were formally adopted in 1999 (Ruffing, 2006). The foundational idea behind the Clusters movement was to better organize content and standards so that students learned academic and technical skills along an integrated continuum. The new Career Cluster framework allowed for more career exploration activities and the ability to expose students to a broader array of career possibilities within their area of interest, rather than a narrow, single-occupation approach to instruction (Ruffing, 2006). The U.S. Secretary of Education at the time of their adoption, Richard Riley, explained, “With this new structure schools can better assure that each student has the opportunity to explore options, set goals and prepare for meaningful work in the new century” (Ruffing, 2006, p. 5). Today, Career Clusters are “an organizing tool for curriculum design and instruction” (Advance CTE, 2016, para. 2) within CTE, and encompass several levels of skills, including foundation, pathway, and essential knowledge and skill statements.

Programs of study. The concept of “programs of study” was unveiled in the 2006 reauthorization of the Carl D. Perkins Career and Technical Education Act. However, this Act only includes a general definition of a program of study. The Act states that programs of study must:

- incorporate secondary education and postsecondary education elements;
- include coherent and rigorous content aligned with challenging academic standards and relevant career and technical content in a coordinated, non-duplicative progression of courses that align secondary education with postsecondary education to adequately prepare students to succeed in postsecondary education;
- may include the opportunity for secondary education students to participate in dual or concurrent enrollment programs or other ways to acquire postsecondary education credits; and
- lead to an industry-recognized credential or certificate at the postsecondary level, or an associate or baccalaureate degree (Association for Career and Technical Education, 2006).

The Department of Education summed up these requirements in an operational definition: “A program of study is a comprehensive, structured approach for delivering academic and career and technical education to prepare students for postsecondary education and career success” (MPR Associates, Inc., 2010, p. 1). Programs of study were envisioned as a singular program construct spanning secondary and postsecondary education, easing student transition between learner levels (MPR Associates, Inc., 2010), and were built on prior initiatives and concepts already in place in CTE, such as School-to-Work and Tech Prep Programs (Klein, et al., 2014). Under the 2006 Perkins Act, local grantees were only required to offer one comprehensive program of study (ACTE, 2006).

Many states and local institutions have embraced the program of study concept though, and are offering more of their CTE content through the program of study organizational construct. In a survey covering the 2008-09 school year, eight secondary CTE state directors

reported that all of their CTE students were enrolled in program of study content (Klein, et al., 2014). However, due to flexibility in the legislation, there is a great deal of variability in the structure and quality of programs of study across the country (Klein, et al., 2014). This led the Department of Education to develop the Program of Study Design Framework, described in detail later in this chapter, as well as to offer grants to aid state development of Rigorous Programs of Study with more standard components in order to assess outcomes (U.S. Department of Education, Office of Career, Technical, and Adult Education, Division of Academic and Technical Education, n.d.b).

Career pathways. While Career Clusters and Programs of Study emerged directly from CTE programs and legislation, the concept of career pathways has much broader applicability. Some states use the terms programs of study and career pathways interchangeably, but for the purposes of this study, career pathways refer to a system of “connected education and training strategies and support services that enable individuals to secure industry relevant certification and obtain employment within an occupational area and to advance to higher levels of future education and employment in that area” (U.S. Departments of Education, Labor, and Health and Human Services, 2012, p. 1). Like programs of study and Career Clusters, career pathways are an organizational framework, but they involve several levels of education and multiple agencies or private sector partners (U.S. Department of Education, Office of Career, Technical, and Adult Education, 2015).

Career pathways emerged from several decades of work in the education, labor, and health and human services arenas (U.S. Department of Education, Office of Career, Technical, and Adult Education, 2015), but are often most associated with adult programs. In recent years, however, the concept has been applied more broadly to a range of audiences and educational

systems, and was included in federal statute in the Workforce Innovation and Opportunity Act of 2014. The new thinking represents a comprehensive, holistic approach to education and workforce systems (U.S. Department of Education, office of Career, Technical and Adult Education, 2015).

Overview of Evaluation

Evaluating the success of program changes and determining results is critical within today's outcome-oriented education environment, especially within CTE. Evaluation can be defined as "the identification, clarification, and application of defensible criteria to determine an evaluation object's value" (Fitzpatrick, Sanders, & Worthen, 2011, p. 7).

Formative vs. summative evaluation. A key component of the evaluation process is identifying the purpose of an evaluation, which is critical to carrying it out in an effective manner. One of the first questions to be asked is whether the evaluation will serve a formative or summative role. Formative evaluation serves the principal purpose of providing information for program improvement (Fitzpatrick, Sanders, & Worthen, 2011). Some examples of formative evaluation in education include collecting information from stakeholders on possible changes to the school calendar, studying the efficacy of implementation of a new teacher peer mentoring program, or developing a feedback program with local business leaders to better tailor curriculum to industry needs.

On the other hand, summative evaluation is designed to aid decisions or judgments about program adoption, continuation, or expansion (Fitzpatrick, Sanders, & Worthen, 2011). It is generally concerned with a program's overall merit in relationship to set criteria. Summative evaluation is more often formal in nature and more often performed by external stakeholders (although neither is required). Examples of summative evaluations include reviewing a CTE

program to determine if it should be continually offered and reviewing a career exploration model to determine if it should be expanded to more schools. A memorable quote by Robert Stake sums up the differences between formative and summative evaluation: “When the cook tastes the soup, that’s formative evaluation; when the guest tastes it, that’s summative evaluation” (cited by Scriven, 1991, p. 19). Generally, both formative and summative evaluation methods are useful in education to meet the goals of program improvement and accountability.

Evaluation approaches. There are numerous evaluation approaches that can be used to measure the success of change in education programs, and multiple ways of organizing and categorizing these approaches. One method of identifying approaches is presented by Fitzpatrick, Sanders, and Worthen (2011), who base their model on the primary factor that directs the evaluation.

The first category of evaluation approaches is oriented to comprehensive judgments of the quality of the program, including expertise-oriented evaluations and consumer-oriented evaluations. They are some of the earliest forms of evaluation and rely heavily on the perceived expertise or influence of the evaluator. The audience is broader than in other evaluations and is often the general public. Accreditation site visits conducted by educational evaluation teams are examples of this approach to evaluation.

The second category of evaluation approaches is oriented to characteristics of the program. This category includes objectives-based, standards-based, and theory-based evaluations. The focus is learning more about some aspect of the program, and logic models are often used to shed light on the underlying rationale behind the program’s activities. These methods of evaluation, focused on objectives, standards, and assessments, now dominate education.

The third category of evaluation approaches are oriented to decisions to be made about the program and include the Context-Input-Process-Product approach, utilization-focused evaluation, as well as other approaches focused on “providing information to improve the quality of decisions made by stakeholders” (Fitzpatrick, Sanders, & Worthen, 2011, p. 123). These evaluations are targeted toward decision makers, such as policymakers, administrators, or program staff, and answer questions specifically designed to aid those individuals in making key program decisions.

Finally, there are evaluation approaches oriented to participation of stakeholders, also known as participatory evaluations. These evaluations focus on involving stakeholders in conducting the evaluation. This can lead to “stakeholders’ greater understanding and ownership of the evaluation leading to greater use of the results or to empowering stakeholders, building evaluation capacity in the organization, and increasing organizational learning and data-based decision making” (Fitzpatrick, Sanders, & Worthen, 2011, p. 189).

Evaluation models. Often, specific evaluation models combine elements of several of these categories, and over time, some models have emerged as leaders in certain areas of education. For example, in 1959, Don Kirkpatrick developed what is now the premier model for evaluating training programs. Primarily designed for use in corporate training settings, the Kirkpatrick Four Levels of Evaluation (Kirkpatrick & Kirkpatrick, 2006) are also highly applicable to education settings as well.

The Kirkpatrick evaluation model consists of four levels: reaction, learning, behavior, and results. Evaluating the first level, reaction, involves determining whether participants react favorably to a learning event, or in other words, whether they enjoyed or were satisfied with the training. The second level, learning, is an evaluation of whether training participants gained the

desired knowledge, skills, or attitudes. The third level, behavior, is a longer-term measurement that entails evaluating whether participants actually apply their learning in authentic situations after training. Finally, the fourth level, results, is defined as whether targeted outcomes, such as increased sales or decreased safety violations, actually occur as a result of the training.

When the levels were first proposed, training evaluation was primarily being conducted at the reaction level. Don Kirkpatrick's model was a major shift in the conversation about the evaluation of training that provided more accurate and complete information about programs' impact using all four levels. Over time, the four evaluation levels have become ingrained in the education and training literature, and have been expanded by other evaluators. For example, Jack Phillips built on the four levels to measure return on investment and also contributed systems and processes to further define them (Kirkpatrick & Kirkpatrick, 2009).

One of the major principles of the Kirkpatrick model is that when designing training programs, you must begin by thinking about the desired end results—level four, and work backward from that point to determine the behaviors that will lead to those results, the knowledge and skills that will improve those behaviors, and the environment that will best relay that information (Kirkpatrick & Kirkpatrick, 2009). This is an especially critical concept to apply to the evaluation of CTE programs, which must maintain a clear focus on desired program outcomes and determining the appropriate conditions to lead to those outcomes.

Evaluation of CTE

In absence of a single set of program quality criteria, there is also no single standard of evaluating CTE programs around the country. There are, however, several national and state-level activities that have had implications for CTE program evaluation in the past and should be

considered in any future efforts. Foundational work to evaluate CTE programs and systems is in place through a number of different initiatives, which will be described in detail below.

Best practices identification. Beginning about a decade ago, there has been an increased focus on the idea of “best practices” in CTE, but this term is often used ambiguously and inconsistently. At their most basic level, best practices are “practices based on research” (Robinson, Shore, & Enersen, 2007, n. p.). However, in many cases, ideas or program components are termed “best practices” with no evaluation evidence to support that claim. While there are numerous research studies linking various CTE interventions or elements to increased student achievement—justifying those elements as evidence-based practices, there is no definitive compilation or source for such information, no comprehensive review of all the best practices in the field.

The National Research Center for Career and Technical Education spent some time exploring this issue and seeking to share more evidence-based practices. In a 2008 presentation, Dr. James Stone, the Center’s director, shared this definition of what he termed a “promising practice” in CTE: “method, technique, program or strategy considered highly successful by practitioners and for which evaluation evidence is available” (Stone & Struebing, 2008, slide 17). The project was renamed “Vanguard Practices in CTE” (NRCCTE, 2012), but unfortunately, reduced funding and administrative issues delayed, and eventually ended, the Center’s work to systematically compile such practices. Principal investigator Oscar Aliaga explains that the project’s goal was to “analyze the different activities, practices, and structures that have led some schools that use CTE as a main component to impact the outcomes of the students” (Aliaga, 2012). While this project yielded some information, its scope was limited and it did not include a focus on the entire CTE system.

Some additional work has been done through the use of systematic reviews of literature, also from the National Research Center, and meta-analysis of CTE studies. Unfortunately, those topics also are limited or too general to offer a complete picture of “best practices” within CTE programs, such as in the case of Valentine et al.’s (2009) work on postsecondary transitions.

Leading to the work of the National Research Center, one of the prior grantees developed a program improvement rubric for secondary-level programs, focused on the question, “What Makes a Career-Technical Education Program Exemplary?” (National Dissemination Center for CTE, 2002). Unfortunately, these criteria, including in Table 2.1, never gained widespread adoption and are no longer in use in any systemic way.

Table 2.1

Criteria from Exemplary Career and Technical Education Programs: Secondary-Level User’s Guide to Program Improvement

Category	Criterion
Program Quality	<ol style="list-style-type: none"> 1. Program goals and objectives are clear, challenging, and appropriate for the intended population. 2. The program content is aligned with learning goals, and is accurate, current, and appropriate for the intended population. 3. The program reflects the vision promoted in recognized academic and occupational standards and by state and national legislation, as appropriate. 4. Collaborations with internal and external organizations and stakeholders are maintained to strengthen the quality and effectiveness of the program.
Educational Significance	<ol style="list-style-type: none"> 1. The program addresses important individual and societal needs. 2. The program contributes to educational excellence for all learners, and leads to other positive results or outcomes. 3. The program design is innovative, reflects current research, and is worthy of replication.
Evidence of Effectiveness and Success	<ol style="list-style-type: none"> 1. The program makes a measurable difference in learning for all participants. 2. The program meets or exceeds identified performance goals. 3. A systematic evaluation process is used to continuously improve the program.
Replicability and Usefulness to Others	<ol style="list-style-type: none"> 1. The whole program or elements of the program can be successfully implemented, adopted, or adapted in other educational settings.

Perkins accountability. The current federal Perkins Act requires local and state programs to report annually on their performance, including on 14 key indicators (U.S. Department of Education, Office of Career, Technical, and Adult Education, Division of Academic and Technical Education, n.d.a). The current indicators are included in Table 2.2. Student-level data to address these indicators is collected by states and then submitted to the federal Department of Education through the state Consolidated Annual Report, which includes the performance data and a narrative summary of the state’s activities under Perkins for the year (U.S. Department of Education, Office of Career, Technical, and Adult Education, Division of Academic and Technical Education, n.d.a). All of the state reports are then compiled into a federally mandated “Report to Congress” on the status of CTE programs around the country. Like many outcomes-only based evaluation products, however, the Report to Congress sheds little light on how to increase CTE program achievement in a systematic way across the country.

Table 2.2

2006 Carl D. Perkins Act Accountability Indicators

Education Level	Indicator
Secondary	1S1: Academic Attainment in Reading/Language Arts 1S2: Academic Attainment in Mathematics 2S1: Technical Skill Attainment 3S1: Secondary School Completion 4S1: Student Graduation Rate 5S1: Secondary Placement 6S1: Nontraditional Participation 6S2: Nontraditional Completion
Postsecondary	1P1: Technical Skill Attainment 2P1: Credential, Certificate or Diploma 3P1: Student Retention or Transfer 4P1: Student Placement 5P1: Nontraditional Participation 5P2: Nontraditional Completion

National Assessment of Career and Technical Education. Also mandated by the Perkins Act, the federal government periodically publishes a comprehensive evaluation of CTE programs around the country. The most recent National Assessment was released in 2014, and included information gleaned from a wide variety of sources, including surveys, case studies, extant data, and commissioned studies on both Perkins implementation and CTE student participation and outcomes (U.S. Department of Education, Office of Planning, Evaluation and Policy Development, Policy and Program Studies Service, 2014). However, this evaluation suffered from a number of difficulties.

When looking at student outcomes, most of the studies are correlational rather than rigorous impact studies designed to assess effects on students. In addition, the report noted that “local subgrantees may offer CTE programs supported by multiple funding sources and may be influenced by other federal, state, and local requirements, making it difficult to disentangle the effects of federal CTE policies from those of other initiatives (U.S. Department of Education, Office of Planning, Evaluation and Policy Development, Policy and Program Studies Service, 2014, p. 13).” The National Assessment also addressed program implementation, but found it difficult to compare programs:

Statutory provisions allow states and local subgrantees considerable flexibility in designing and implementing POS. This flexibility has led to variation — both across and within states — in how CTE practitioners define key terms, design programs, and organize, administer, and deliver services (U.S. Department of Education, Office of Planning, Evaluation and Policy Development, Policy and Program Studies Service, 2014, p. 54).

The prior National Assessment, funded by the previous iteration of Perkins, was released in 2004. Its key findings were similar, as the authors stated “In large part, the pace and path of improvement are hampered by a lack of clarity over the program’s fundamental purpose and goal” (Silverberg, Warner, Fong, & Goodwin, 2004, p. xviii). The findings from these two studies lend credence to the idea that a more systematic national evaluation system that can clearly focus on individual programs, their elements and their impact, would be helpful.

State efforts. In addition to mandating accountability indicators and the National Assessment as described above, the Perkins Act requires states to undergo evaluation-related activities in a number of areas, including approving recipients for funding, annually evaluating the effectiveness of CTE programs, and measuring the progress of students (Association for Career and Technical Education, 2006). States have chosen to implement these requirements in a variety of ways, using accountability data reviews, local applications, and program surveys (Lintner, 2013).

Some states have gone even further to provide in-depth quality review instruments to be used in local programs. For example, Missouri has developed a set of Common Criteria and Quality Indicators, “designed to provide guidance and direction to local school districts in establishing, maintaining, and evaluating quality career education programs” (Missouri Department of Elementary and Secondary Education, Office of College and Career Readiness, 2015, p. 2). These criteria are accompanied by a program evaluation and improvement tool that can be used by local educators to evaluate their programs. A number of additional states have similar tools, designed either for program approval, program improvement, or recognition. Some of these tools are designed to be self-assessments, and others are designed to be administered by the state, and they differ significantly in level of detail and elements included.

National frameworks. Outside of the Perkins Act requirements, a number of nationally-known organizations have sought to fill gaps in program evaluation by developing their own set of program quality criteria, unique to the program models they each promote.

The Department of Education has sought to address the need for more standard program quality elements by building on the provision in the 2006 Perkins Act that requires states and local school districts to develop programs of study. Officials at the Department expanded on the basic description of a program of study included in the Act by compiling a more explicit set of quality elements required for such programs through a large-scale technical assistance project. The resulting Program of Study Design Framework (RPOS Framework) was unveiled in 2010, and contains 10 components and related subcomponents that the federal Office of Vocational and Adult Education (OVAE), now known as the Office of Career, Technical and Adult Education (OCTAE), suggested “support the development and implementation of effective programs of study” within CTE (OVAE, 2010, p. 1). The intent of the Department of Education was that the RPOS Framework “provide a useful quality assurance marker for states seeking to promote local development of consistent POS that comply with congressional intent and promote program improvement within Perkins-funded programs” (MPR Associates, Inc., 2010, p. 6). The RPOS Framework’s components and subcomponents are listed in Table 2.3.

Since the release of the Programs of Study Framework, there have been numerous national efforts to spread its use. The Department of Education has sponsored national grant competitions to support state and local efforts to develop programs of study using the Framework, used it as the basis for its technical assistance efforts, and funded the development of a local capacity assessment based on the Framework. The local capacity self-assessment is primarily designed to aid states in “providing technical assistance to local program of study

Table 2.3

Program of Study Framework Components and Subcomponents

Component	Subcomponents
1. Legislation and Policies	<ul style="list-style-type: none"> • Provide for state and/or local funding and other resources, such as professional development and dedicated staff time, for POS development. • Establish formal procedures for the design, implementation, and continuous improvement of POS. • Ensure opportunities for any secondary student to participate in a POS. • Require secondary students to develop an individual graduation or career plan. • Provide resources for long term sustainability of POS.
2. Partnerships	<ul style="list-style-type: none"> • Create written memoranda of understanding that elaborate the roles and responsibilities of partnership members. • Conduct ongoing analyses of economic and workforce trends to identify statewide (or regional) POS to be created, expanded, or discontinued. • Link into existing initiatives that promote workforce and economic development, such as sector strategies and other activities supported by the Workforce Investment Act. • Identify, validate, and keep current the technical and workforce readiness skills that should be taught within a POS.
3. Professional Development	<ul style="list-style-type: none"> • Support the alignment of curriculum from grade to grade (9-12) and from secondary to postsecondary education (vertical curriculum alignment). • Support the development of integrated academic and career and technical curriculum and instruction (horizontal curriculum alignment). • Ensure that teachers and faculty have the content knowledge to align and integrate curriculum and instruction. • Foster innovative teaching and learning strategies (see #9 below).
4. Accountability and Evaluation Systems	<ul style="list-style-type: none"> • Include the “10 Essential Elements of a State Longitudinal Data System” identified by the Data Quality Campaign. • Provide for administrative record matching of student education and employment data (i.e., Unemployment Insurance (UI) wage records). • Yield valid and reliable data on key student outcomes (indicators) referenced in Perkins and other relevant federal and state legislation. • Provide timely data to evaluate and improve the effectiveness of POS.
5. College and Career Readiness Standards	<ul style="list-style-type: none"> • Be developed and continually validated in collaboration with secondary, postsecondary, and industry partners. • Incorporate essential knowledge and skills (i.e., academic skills, communication, and problem-solving), which students must master regardless of their chosen career area or POS. • Provide the same rigorous knowledge and skills in English and mathematics that employers and colleges expect of high school graduates. • Incorporate industry-recognized technical standards that are valued in the workplace. • To the extent practicable, be internationally benchmarked so that all students are prepared to succeed in a global economy.

Table 2.3 (Continued)

6. Course Sequences	<ul style="list-style-type: none"> • Map out the recommended academic and career and technical courses in each POS. • Begin with introductory courses at the secondary level that teach broad foundational knowledge and skills that are common across all POS. • Progress to more occupationally-specific courses at the postsecondary level that provide knowledge and skills required for entry into and advancement in a chosen POS. • Offer opportunities for students to earn postsecondary credit for coursework taken during high school.
7. Credit Transfer Agreements	<ul style="list-style-type: none"> • Provide a systematic, seamless process for students to earn college credit for postsecondary courses taken in high school, transfer high school credit to any two- and four-year institution in the state that offers the POS, and transfer credit earned at a two-year college to any other two- or four-year institution in the state that offers the POS. • College credit should be automatically transcribed at the college for high school students so that they can transfer seamlessly into the postsecondary portion of a POS without the need for additional paperwork or petitioning for credit. • Describe the expectations and requirements for, at a minimum, teacher and faculty qualifications, course prerequisites, postsecondary entry requirements, location of courses, tuition reimbursement, and credit transfer process.
8. Guidance Counseling and Academic Advisement	<ul style="list-style-type: none"> • Are based on state and/or local guidance and counseling standards, such as the National Career Development Guidelines. • Ensure that guidance, counseling, and advisement professionals have access to up-to-date information about POS offerings to aid students in their decision making. • Offer information and tools to help students learn about postsecondary education a career options, including prerequisites for particular POS. • Offer resources for students to identify their career interests and aptitudes and to select appropriate POS. • Provide information and resources for parents to help their children prepare for college and careers, including workshops on college and financial aid applications. • Offer Web-based resources and tools for obtaining student financial assistance.
9. Teaching and Learning Strategies	<ul style="list-style-type: none"> • Be jointly led by interdisciplinary teaching teams of academic and career and technical teachers or faculty. • Employ contextualized work-based, project-based, and problem-based learning approaches. • Incorporate team-building, critical thinking, problem-solving, communication skills, such as through the use of career and technical student organization (CTSO) activities.
10. Technical Skills Assessments	<ul style="list-style-type: none"> • Measure student attainment of technical skill proficiencies at multiple points during a POS. • Employ industry-approved technical skill assessments based on industry standards, where available and appropriate. • Employ State-developed and/or approved assessments, particularly where industry-approved assessments do not exist. • Result in the awarding of secondary credit, postsecondary credit, or a special designation on a student’s high school diploma. • Incorporate performance-based assessment items, to the greatest extent possible, where students must demonstrate the application of their knowledge and skills.

developers and reviewing state CTE program approval processes,” (MPR Associates, Inc., 2010, p. 6), but may be applicable for other purposes as well. However, the Framework has not been formally reviewed, and its elements have not been tested for validity.

Two other evaluation models have gotten widespread attention at the national level. With the rise in attention on the career academy model within CTE, the National Career Academy Coalition convened a group of experts to develop a set of National Standards of Practice that were released in 2004. These standards focus on (1) defined mission and goals, (2) academy structure, (3) host district and high school, (4) faculty and staff, (5) professional development, (6) governance and leadership, (7) curriculum and instruction, (8) employer, higher education, and community involvement, (9) student assessment, and (10) cycle of improvement (National Career Academy Coalition, 2013).

Experts in the Linked Learning field, a CTE model developed by ConnectEd in California, have also developed a set of criteria to guide program quality reviews and continuous improvement. Criteria are divided into four main areas as outlined in Table 2.4 (Linked Learning, 2012).

Table 2.4

Criteria for Linked Learning Pathway Quality Review and Continuous Improvement

Pathway Design	Engaged Learning	System Support	Data and Impact
1.1 Design Structure	2.1 Standards-aligned Curriculum	3.1 District Policies	4.1 Pathway Evaluation
1.2 Governance	2.2 College and Career Readiness	3.2 Leadership	4.2 Student Data
	2.3 Real-world Relevance	3.3 Professional Development	
	2.4 Integrated Curriculum	3.4 Qualified Staff	
	2.5 Instruction and Assessment	3.5 Partnerships	
	2.6 Work-based Learning		
	2.7 Personalization and Support Services		

The program quality criteria developed by the National Career Academy Coalition and Linked Learning and validated by experts in those fields begin to shed light on important elements of CTE programs, but are specific to those program models. CTE is a very diverse enterprise, and efforts to define quality in a way that is more broadly applicable must take a step back from specific models and focus on the basic elements that are applicable to all CTE programs, regardless of delivery mechanism.

ACTE High-quality CTE Program of Study Research

It was with the goal of trying to identify those basic elements that were applicable to all CTE programs that ACTE undertook its high-quality CTE initiative. This initiative is designed “as a multi-step project to identify a comprehensive, research-based quality CTE program of study framework, test the framework, and integrate it into our efforts to recognize and disseminate information on best practices within CTE” (Imperatore & Hyslop, 2015a, p.1). The first phase of the project, its “defining” phase, has already been completed through prior research. This study is situated as a key component of the second phase of the project, “testing the framework,” and builds extensively on the initial framework research. Due to the significant relationship between this study and ACTE’s earlier work, the previous research contained in Hyslop & Imperatore’s (2015a) *Defining High-quality CTE: Contemporary Perspectives on CTE Quality*, *Defining High-Quality CTE: Quality CTE Program of Study Framework, Draft Version 1.0* (Imperatore & Hyslop, 2015b), *Defining High-Quality CTE: Quality CTE Program of Study Framework, Draft Version 3.0* (Imperatore & Hyslop, 2016), and *Defining High-Quality CTE: Quality CTE Program of Study Framework, Beta Version 4.0* (Imperatore & Hyslop, 2017) is summarized extensively below.

Analysis of current quality frameworks. ACTE’s research began by seeking to identify current sources of information about high-quality CTE—documents and statements that were already being used in the field and had implications for CTE quality. Three types of documents were identified: broad national organization positions, state documents, and national frameworks. In general, the national organization positions came in the form of speeches, white papers or shorter documents that included only general principles of CTE quality. Very few specific details were included in this document, and while the rhetoric provided some insights into the national conversation, these documents were generally determined to be too broad to have implications for a more detailed framework.

The second category identified was state documents, like those described in the prior section. At least some policy, evaluation tool, or program improvement resource has been developed in almost every state in absence of a single national program quality standard. Finally, ACTE identified more detailed national frameworks. In addition to the frameworks identified in the prior section, like the RPOS Framework and the Career Academy National Standards of practice, this category included a range of targets, including career pathways, institutions and individual programs. Partnering with REL Central, the federally funded Regional Education Lab in Denver, Colorado, these last two sets of documents were divided, with REL Central undertaking a review of a sample of state policy documents, while ACTE began its in-depth analysis of the national frameworks.

Using a content analysis process, ACTE sought to answer how and why the frameworks had been developed, how they were structured, and what content elements were included. The first step in the process was to narrow the sample for analysis. By consulting experts, conducting online searches, and carefully reviewing documents, 21 frameworks were identified for inclusion

in the study. These frameworks were further divided into system-level frameworks, which were applicable to multiple CTE programs or pathways across a geographic area, or documents that focused on a broad range of populations or partners; institution-level frameworks, which covered all the CTE-related programs or elements across an entire institution (generally a high school or community college); and program-level frameworks, which focused on single CTE programs. By far, the largest category was the systems-level frameworks, which included 10 documents.

To answer their first two research questions, ACTE researchers examined the purpose and structure of each of the documents, using supplemental information from publically available sources where necessary, such as the websites that housed the documents. Some of the key findings included that many of the frameworks lacked a defined research base, and instead relied solely on expert validation; that there was great variability in the structure, scope and level of detail of the documents; and that the majority of the frameworks were designed for formative or a mix of formative and summative evaluation. Most of the frameworks provided few specifications, which could impact the utility of the documents for practitioners. Several of the frameworks did include additional rubrics or supplemental information though, which added to the detail provided.

In order to answer the last research question about the content of the documents, a set of content codes was developed by REL Central and revised by ACTE researchers to fit the national frameworks. A total of 741 pieces of content were coded across the frameworks using 68 codes under the following "master categories":

- legislation/state policies
- program/school/system characteristics
- teacher-related elements

- assessment
- instructional delivery
- program content
- student recruitment and support
- aspects of career exploration and guidance
- partnership characteristics
- types and use of data
- equipment and space concerns
- progress toward framework goals

The most common codes across the frameworks reflected many of the conversations going on across CTE and education more generally, with the top 3 codes being the involvement of business partners, the involvement of education partners, and the use of accountability and outcomes data. The least commonly used codes tended to be much narrower, like the formal recognition of prior learning and the ability of a program to scale progress.

The examination of structure, purpose and content of the currently available national frameworks led the researchers to the following conclusions:

- "The most commonly found content elements are consistent across many of the frameworks, and align with other initiatives in CTE" (Imperatore & Hyslop, 2015a, p. 11).
- "The diversity of scope and structure, along with the consistent lack of detail...does leave room for greater clarity and consistency in the conversation about high-quality CTE, particularly as it relates to individual local programs" (Imperatore & Hyslop, 2015a, p. 11).

These conclusions led the researchers to suggest ACTE's work focus on the program of study as the unit of analysis, as this was the biggest gap identified in the currently available frameworks. The results of the content and structural analysis served to inform the next stage of ACTE's research process involving the drafting of a comprehensive framework.

Development of ACTE Framework. Work on the development of the actual ACTE Framework began in August 2015, and the first draft was released in November 2015 as *Defining High-Quality CTE: Quality CTE Program of Study Framework, Draft Version 1.0* (Imperatore & Hyslop, 2015b). To develop the Framework, researchers began by grouping the identified content codes into a number of main ideas, and pulling specific content from the documents covered in the original research to identify key descriptors around each of the codes. In addition, content from state documents and from several documents that were excluded from the original research because they were too narrow were examined as relevant to specific codes. Both researchers worked to identify the criteria, and then discussed and came to consensus on specific language. Through this process, 12 main elements emerged, with a total of 105 criteria. The original 12 elements included:

- standards-aligned and integrated curriculum
- sequencing and articulation
- student assessment
- prepared and effective program staff
- engaging instructional strategies
- access and equity
- facilities and equipment
- business and community partnerships

- career development
- career and technical student organizations
- work-based learning
- data and outcomes

The first draft of the Framework was released at the Association for Career and Technical Education's CareerTech VISION 2015 conference. At that event, two focus groups and a more general feedback session were held around the Framework criteria. Additional focus groups were held in early 2016 with both state and local CTE administrators, and informal feedback was sought in a number of forums. The Framework was then revised slightly, and a second version was shared to solicit more detailed opinions. First, individuals that had participated in one of the earlier feedback sessions were sent a detailed survey to complete. In addition, key experts, both within and outside the CTE community, were identified in relation to each of the 12 main Framework elements. A total of 30 experts reviewed relevant pieces of the Framework and provided feedback. Responses from the survey and expert review were incorporated into *Defining High-Quality CTE: Quality CTE Program of Study Framework, Draft Version 3.0* (Imperatore & Hyslop, 2016), which was publically released in November 2016.

A few questions about the elements and criteria remained following the expert review, and these were addressed in several ways, including another feedback session of ACTE members, additional literature review on the items in question, and follow-up conversations with experts. Finally, several more edits were made to the Framework to standardize language throughout and ensure consistency across elements. The resulting Framework *Beta Version 4.0* was used in the pilot study and its elements can be found in their entirety in Appendix D (as part of the evaluation instrument).

To transition the Framework elements and criteria into a full evaluation instrument, researchers reviewed rubrics obtained during earlier stages of the research project, as well as prior research on evaluation scales. A four-part scale was developed to apply to each criterion, ranging from “0,” representing criteria not being achieved at all, to “3,” representing criteria that are fully implemented throughout the entire program of study. Corresponding descriptions to elaborate on each rating were drafted and shared during a session of CTE data and program experts, then revised to reflect feedback. The rating scale can be found in Appendix C. Both the rating scale and the Framework elements and criteria were then incorporated into an online evaluation instrument, and it was this resulting instrument that was the focus of the pilot testing in this study. A draft of the entire evaluation instrument can be found in Appendix D.

Conceptual Framework

The constructs of program evaluation and instrument development formed the basis of the conceptual framework for this study. As stated earlier, evaluation can be defined as “the identification, clarification, and application of defensible criteria to determine an evaluation object’s value” (Fitzpatrick, Sanders, & Worthen, 2011, p. 7). ACTE’s early research on which this study was based set out to identify the “defensible criteria” that could be used to systematically identify and replicate quality within CTE programs of study, and then design an appropriate evaluation process.

Program theory-driven evaluation. The idea of determining appropriate conditions leading to desired educational outcomes provided a key element of the theoretical foundation for this study. Theory-driven evaluations, part of the category of evaluation approaches oriented to characteristics of a program, are based on program theory. They focus on “investigating the causal mechanisms by which a program achieves its effects” (Davidson, 2000, p. 18). At its

most basic level, “program theory is focused on the ‘main active ingredients’ that are presumed to lead to desired outcomes, and the key conditions under which they are believed to operate” (Donaldson, 2007, p. 23). One of the simplest models of program theory is illustrated in Figure 2. While actual program theories are generally much more complex, the underlying concept of a program’s elements leading to its outcomes is maintained even as program complexity grows (Coryn, Noakes, Westine, & Schroter, 2010).

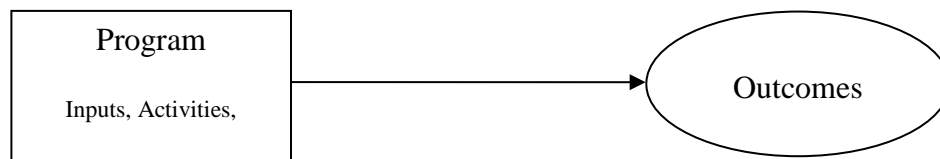


Figure 2.1. Linear program theory model. Adapted from Program Theory-driven Evaluation Science: Strategies and Applications, by S. I. Donaldson, 2007, New York, NY: Routledge.

Theory-driven evaluation is often aligned with the phenomenon of “unpacking the black box,” first proposed by Edward Suchman in 1967 to emphasize the importance of evaluating not only the overall effectiveness of social programs, but understanding the components of such programs. Donaldson (2007) defines a three-step process for implementing theory-driven evaluation. First, an understanding of how a program is hypothesized to achieve its outcomes must be developed. Second, evaluation questions are then formulated, and then finally, answered.

This process parallels the proposed purpose of this study—to validate the practices or “components” that will lead to improved student outcomes in CTE. Coryn, Noakes, Westine,

and Schroter (2010) assert that such a theoretical approach will result in information that is more useful to policymakers, and Rogers (2000) claims that theory-driven evaluation “can lead to better information about a program that is important for replication or for improvement” (p. 232). These benefits directly address the underlying concerns leading to this study.

Validity of evaluation instruments. Gall, Gall, and Borg (2007) present that the validity of an evaluation instrument “refers to the appropriateness, meaningfulness, and usefulness of specific inferences made from test scores” (p. 151). This premise is based on standards for validity evidence set by the American Educational Research Association (AERA), the American Psychological Association (APA), the National Council on Measurement in Education (NCME), and the Joint Committee on Standards for Educational and Psychological Testing in *Standards for Educational and Psychological Testing* (2014). In those standards, validity “refers to the degree to which evidence and theory support the interpretation of test scores for proposed uses of tests” (p. 11). The authors of the standards emphasize that it is the interpretations or inferences made using the results of the instrument that are valid or invalid, not the instrument itself, and that evidence of this validity is the “most fundamental consideration in developing tests and evaluating tests” and a “necessary condition for the justifiable use of the test” (p. 11).

Trochim (2006) focuses on concepts highlighted in earlier versions of the standards to suggest that it is critical to focus on the validity of the “operationalization” of a measure, and elaborates: “Any time you translate a concept or construct into a functioning and operating reality (the operationalization), you need to be concerned about how well you did the translation” (para. 1). He presents the term “construct validity” to represent the overarching idea of ensuring that any operationalization of an idea accurately reflects the underlying construct it is attempting

to measure (in this case, high-quality CTE). This makes construct validity particularly important as this research seeks to determine whether the evaluation instrument for high-quality CTE can actually be used to infer program quality. There are two primary ways to ensure construct validity—translation validity and criterion-related validity (Trochim, 2006).

Translation validity. Translation validity played a key role in the first step of the theory-driven evaluation process and thus, the first step in the development of the ACTE Framework and evaluation instruments. During the initial phase of its project, researchers subjected the ACTE Framework and evaluation instrument to several tests of translation validity, including face validity and content validity. According to Drost (2011), “Translation validity attempts to assess the degree to which constructs are accurately translated into the operationalisation, using subjective judgment—face validity—and examining content domain— content validity” (p. 116).

Face validity, as its name implies involves reviewing the operationalization to see if “on its face” it appears to represent the construct. Alone, this type of validity measure is by definition subjective, but can be improved with intentionality, such as through the use of a group of experts to review (Trochim, 2006). Content validity involves comparing the operationalization with the relevant content domain of the construct—essentially the criteria that define a program (Trochim, 2006). This can be difficult in social science research because often no consensus exists on construct definitions, but can be accomplished in two ways: getting the opinion of experts in the field, or asking a number of questions about an instrument’s alignment (Drost, 2011). To address both of these types of validity, content analysis of other program quality evaluation tools and expert review and input were incorporated into ACTE’s Framework design.

Criterion-related validity. Drost (2011) defined criterion-related validity as “the degree of correspondence between a test measure and one or more external referents (criteria), usually measured by their correlation” (p. 118). The research questions of this study were based on the principles of criterion-related validity, particularly focused on establishing the predictive validity of the instrument. This relates directly to the concept of “validity evidence based on relationships to other variables” as presented in the *Standards for Educational and Psychological Testing* (AERA, et al., 2014), which outlines the process of relating underlying constructs to external variables to provide evidence of validity in test development.

One of the major principles of a leading evaluation model by Kirkpatrick and Kirkpatrick (2009) is that when designing training programs, you must begin by thinking about the desired end results—level four, and work backward from that point to determine the behaviors that will lead to those results, the knowledge and skills that will improve those behaviors, and the environment that will best relay that information. This is an especially critical concept to apply to the evaluation of CTE programs, which must maintain a clear focus on desired program outcomes and determining the appropriate conditions to lead to those outcomes.

To recap, based on this conceptual framework focused on ACTE’s early research, program-theory evaluation and principles of validity, it was expected that the elements that were identified and validated on their constructs in the initial stage of ACTE’s research would lead to desired program outcomes measured by student performance. CTE programs that score higher on the evaluation instrument will have higher levels of student achievement, and programs that score lower overall on the evaluation instrument will have lower overall outcomes on student achievement measures.

Summary of Literature Review

Evaluation can be defined as “the identification, clarification, and application of defensible criteria to determine an evaluation object’s value” (Fitzpatrick, Sanders, & Worthen, 2011, p. 7). CTE programs around the country have a long history of federal investment and growing support, but are often criticized for inconsistent quality and an inability to identify and replicate best practices. While federal legislation provides a core set of outcomes indicators, these outcome indicators are not enough to truly evaluate CTE to determine its value—or how to increase that value. In absence of a standard set of national criteria for determining CTE’s worth, numerous organizations, as well as states, have stepped in to develop tools and resources to fill gaps in the CTE program evaluation infrastructure. However, many of these tools and resources are not applicable broadly at the local CTE program level.

It was this gap that ACTE set out to fill through the development of its high-quality CTE Framework and accompanying self-assessment tool. The conceptual framework of program-theory driven evaluation, including the idea that necessary inputs will lead to desired outcomes (Donaldson, 2007), guided this work and the current study. Principles of validity were applied throughout, driving the development and testing of the Framework.

CHAPTER 3: METHODOLOGY

The purpose of this study was to build on the research already completed by ACTE in the development of the ACTE Framework and corresponding evaluation instrument to examine the validity of the instrument's results and its utility when used independently by local educators in a pilot test. In this study, the instrument was tested particularly for its criterion validity—whether the included elements are appropriate measures of program quality, as correlated to already established program outcome measures—as well as its utility, including ease of use, practicality, and usefulness of results to local practitioners. The specific research questions that were examined to achieve the purposes of the study are:

1. Does the program evaluation instrument provide valid evidence of program quality?
2. Is the evaluation instrument a useful tool for local educators to assess the quality of their CTE programs?

This chapter is divided into three sections. The first section presents the two-phase research design that was used to achieve the purpose of the study. The second and third sections outline the research procedures, including sampling, instrumentation, data collection and data analysis, used in each phase of the research.

Research Design

This research study was conducted as a two-phase, sequential mixed-method study, grounded on an evaluation instrument measuring the ACTE Framework for program quality and the premises of theory-driven evaluation. A mixed-method design was selected to capitalize on the strengths of both qualitative and quantitative approaches and answer the very distinct

questions posed in the study related to validity and utility. A summary of the research design, based on the presentation of sequential, mixed model research in Tashakkori and Teddlie (2003) is shown in Figure 3.1.

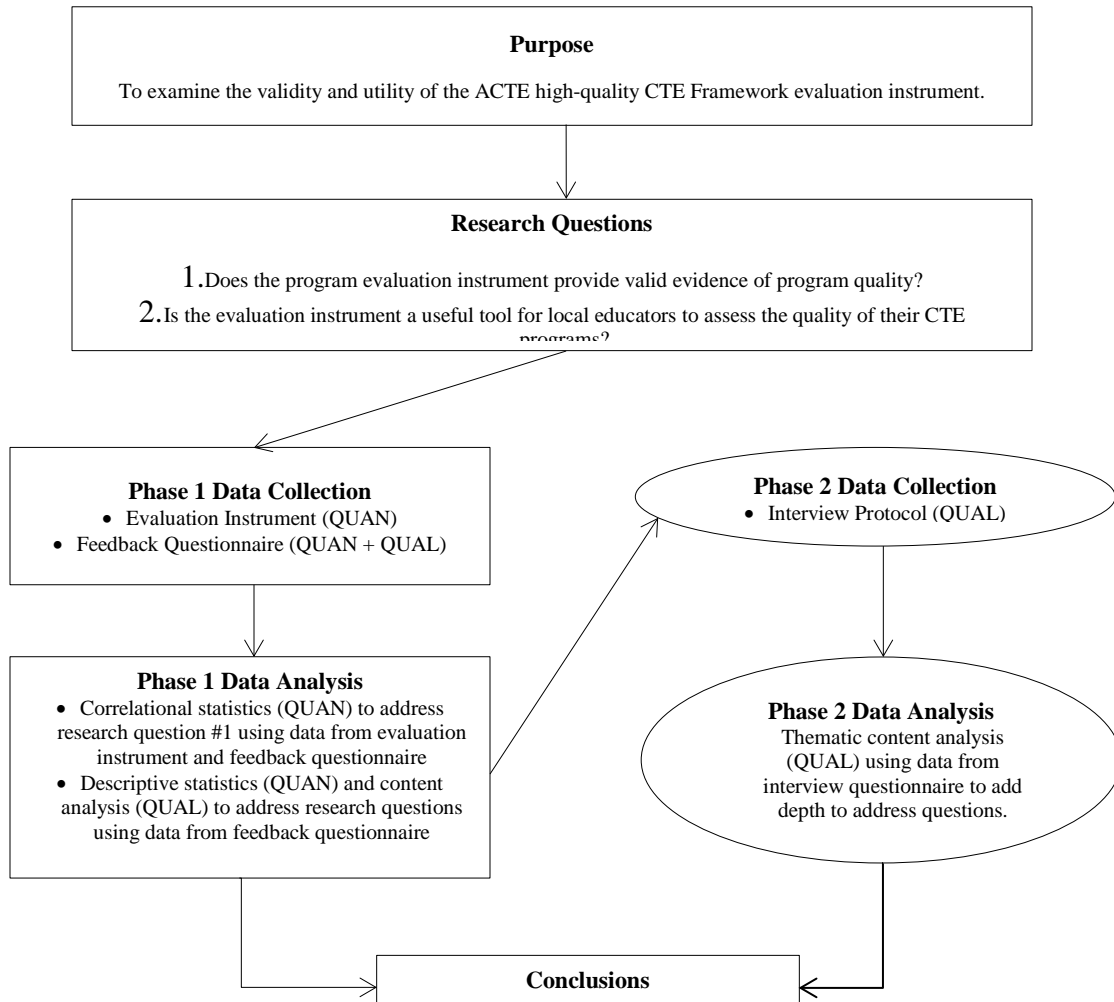


Figure 3.1. Two-phase, sequential mixed methods research design.

The first phase of this study included a correlational research component to determine the criterion validity of the evaluation instrument by measuring the extent to which responses on the evaluation instrument were associated with actual indicators of program performance, and a

feedback questionnaire targeting qualitative and quantitative data to capture the experience of participants as they completed the evaluation. Locke, Spirduso, and Silverman suggest that such a blending of qualitative and quantitative evidence can reinforce findings and enrich conclusions (2007), and this was particularly useful when addressing the questions of utility for wide-scale use addressed by this study. As such, during the first phase of the project, both qualitative and quantitative data were collected concurrently, with the principles of triangulation applied as suggested by Terrell (2012) in his “concurrent triangulation strategy,” where data are collected from different methods and integrated for “confirmation, corroboration or cross-validation” (p. 268).

This was followed by the second phase of the research project, a qualitative research component involving interviews with a select group of Phase One participants to gather more in-depth perspectives. Based on the sequential, mixed model design as described by Tashakkori and Teddlie (2003), the second phase of the project built on the findings and inferences from the first phase of the project. The small set of in-depth interviews was used to confirm and explain initial findings, as well as add depth and detail. This interview phase of the project yielded only qualitative data, which was then combined with phase one data and inferences to provide more comprehensive answers to the study’s research questions.

Phase One

Program sample. The goal for the study was a sample of a minimum of 30 individual CTE programs of study, as defined in Chapter 1. Briefly, a program of study consists of a clearly defined sequence of academic and technical courses spanning secondary and postsecondary education and leading to a credential (generally either a postsecondary certificate or degree or industry certification). While ideally a program of study would be jointly developed

and offered by secondary and postsecondary institutions, that level of integration was not be required for this study—a program could have been concentrated at the secondary or postsecondary level.

Each program was represented in the study by a key contact, either a CTE teacher or administrator, who agreed to participate in the research and complete required data collection instruments. A single individual could submit multiple program “participants” to the study. Gall, Gall, and Borg (2007) suggest that 30 is the minimum sample size needed for valid correlational research, which is the statistical approach to be performed on the quantitative data for the first research question. Unfortunately, one piece of the data collection—the Perkins performance data described below—proved to be much more difficult than expected, and while 39 programs participated in the study, only 23 of those were able to provide this data for the correlational research, reducing the power of those statistical results (Grimm, 1993).

Participants were selected through a non-probability convenience sample, with volunteers considered first. Throughout the foundational research that led to this study, a list of potential volunteers was collected. These individuals were contacted and given the first opportunity for their programs to participate in the study. While a convenience sample is limited in its ability to be generalized (Gall, Gall, & Borg, 2007; Trochim, 2006), a complete sampling frame of the entire population of CTE programs of study does not exist, and the time and effort required for participation in this study required a vested interest.

After reaching out to the list of compiled volunteers, additional information was circulated broadly throughout the entire CTE community in an attempt to obtain more participants. Information was distributed by e-blast to approximately 20,000 ACTE members on multiple occasions, and specific communications were sent to the organization’s leadership and

volunteer groups, such as the Board of Directors, state association leadership and “Educators-in-Action.” Presentations were also held at several national and state-level conferences to explain the project and purposefully solicit volunteers.

To aid with the issue of generalizability, purposive sampling techniques were also used to achieve geographic diversity and diversity of CTE program areas, using the membership of ACTE as the sampling frame. Gall, Gall and Borg (2007) suggest that “inferential statistics can be used with data collected from a convenience sample if the sample is carefully conceptualized to represent a particular population” (p. 176). Direct communications were sent to specific members or requesting specific program areas to complete the sample in an attempt to represent as much of the diversity within CTE programs of study across the country as possible.

Data. The purpose of the study calls for the examination of the validity of the instrument’s results and its utility. The validity of an instrument is generally referred to as “the appropriateness, meaningfulness, and usefulness of specific inferences made from test scores” (Gall, Gall, & Borg, 2007, p. 151). In this case, to examine *criterion validity*, the following indicators were used: program quality and program performance. Program quality was measured by the overall score on ACTE’s high-quality CTE Framework evaluation instrument, and program performance was measured using participant’s federal CTE data as reported under the Carl D. Perkins Career and Technical Education Act’s core indicators of performance. To determine the validity of the instrument’s results as a representation of quality, a separate correlation was estimated between the score on the evaluation instrument measuring elements of program quality and each of three indicators of program performance: technical skill attainment, completion, and placement. Additional evidence of the instrument’s validity was collected through the feedback questionnaire to supplement the correlation.

It is important to note that the Perkins performance data, in addition to being very difficult for some programs to obtain, also contain potential limitations. Different states define CTE students and programs differently, so data do not always represent the same underlying population or exposure to CTE. In addition, data reporting could lag a year or more, depending on the timing of state data collection and feedback. This means that data may not be reflective of the current state of programs. Despite these limitations, it is the only current information available on CTE program performance that has been standardized in any way, and the three measures chosen are generally the most complete and relevant data available.

In turn, *instrument utility* refers to its ease of use, practicality, and usefulness of results to local practitioners. For the purpose of this study, instrument utility was determined through participants' perspectives as collected through both open- and closed-ended questions to be answered immediately upon their completion of the evaluation instrument on a feedback questionnaire.

Instrumentation. To gather data on validity and utility of the evaluation instrument in the first phase of the study, the ACTE Framework evaluation instrument was used along with a feedback questionnaire. The development of the evaluation instrument was summarized in the review of literature, and can be found in Appendix D. The evaluation instrument was the result of extensive research and evaluation by ACTE in the early stages of its high-quality CTE initiative (Imperatore & Hyslop, 2015a; Imperatore & Hyslop, 2015b; Imperatore & Hyslop, 2016; Imperatore & Hyslop, 2017). It consists of the entire set of 12 elements and 102 criteria from the ACTE Framework, combined with a rating scale, and can be found in Appendix D. The Framework elements represent broad concepts on which high-quality CTE programs should be built, ranging from standards-aligned and integrated curriculum to work-based learning. Each of

the elements is then further specified through a number of criteria outlining key aspects of that element. The rating scale asked participants to rate each individual criterion with a rubric using a scale of zero to three, with zero representing no evidence of the criterion in the program of study, and three representing substantial achievement of the criterion in the program of study. The total score possible on the rubric ranged from 0 to 306. As outlined in Chapter 2, the evaluation instrument was made available online through Survey Monkey for ease of participation.

The second piece of instrumentation for the first phase of the project was a short feedback questionnaire that was completed by participants in conjunction with the evaluation instrument, to gather perspectives on the process participants engaged in to complete the rubric, their overall thoughts on its utility, and the program performance data necessary for the correlation with scores on the evaluation instrument. It included both open-ended questions and questions asking for numerical responses that yielded quantitative data. Examples of questions included the following:

- How well did you think the results you reported on the evaluation instrument accurately reflected the quality of your program?
- How would you rate the overall ease of use of the evaluation instrument?
- Are there any changes that should be made to the tool to allow it to be used by more programs?

The feedback questionnaire was developed by the researcher, in conjunction with research partners at ACTE, and was shared with a potential group of study participants and education researchers to provide expert review and validation. Feedback from this validation was used to make necessary changes to the questionnaire, and the final version is included in

Appendix E. This feedback questionnaire was also made available through Survey Monkey for ease of participant access and data collection.

Data collection procedures. Potential participants were contacted by email and provided detailed information about the study, including the purposes of the research, background on ACTE's high-quality CTE initiative, the research process and instruments, expected time required for participation, procedures to be followed, and potential benefits to participating. Participants were also assured of the voluntary nature of the study. If participants expressed interest, they were asked to consent to participation by clicking on the link to the online Survey Monkey evaluation instrument and feedback questionnaire. Participants were also provided with a PDF of the evaluation instrument and instructed that they may wish to fill out the form on paper first before transferring information into Survey Monkey to ensure they had the time necessary to complete.

Reminders were sent regularly to check the progress of those who were interested in participating. Data were collected on a rolling basis from late December 2017 through August 2018, when it was determined that we had reached market saturation of those able to participate and further efforts would not yield significant additional participants. Significant time was spent communicating with participants and following up on missing research elements. Many individuals did not initially submit both instruments, or did not include their Perkins performance data as required on the feedback questionnaire. The majority of those with missing Perkins data that were ultimately able to obtain it forwarded this data via email to be added to their input manually.

Data analysis. Before analysis could begin, data from the evaluation instrument were organized and linked with data from the feedback questionnaire's Perkins performance measures.

In addition, descriptive statistics on the participants' performance on the evaluation instrument were calculated, along with statistics related to the internal structure of the instrument, including internal consistency. Additional analysis of data collected in Phase One of the study is described below for each of the research questions.

Question one. To address the first research question, "Does the program evaluation instrument provide valid evidence of program quality?" quantitative data from the completion of the evaluation rubric and program performance data from the feedback questionnaire were used. Each participant's scale score (0 to 302) on the evaluation instrument was correlated with each of the performance levels for the program's Perkins performance indicators, as reported on the feedback questionnaire, to determine the degree of the relationship between scores on the program quality framework and program outcomes.

The Pearson Product Moment Correlation statistic, r , was used to measure the correlation since both variables are measured by at least interval level data as required (Trochim, 2006). The statistical program SAS was used to compute a correlation coefficient for the quality scale and each individual performance level, resulting in three different measures—one correlation each for technical skill attainment, completion and placement. Shared variance, or r^2 , was also calculated. Due to the small sample size, alpha was set at $p < .10$ to reduce the Type II error and help compensate for the lack of power. Grimm (1993) suggests this more liberal value of alpha can be appropriate for new or exploratory areas of research in which the consequences of mistakenly rejecting the null hypothesis are not extremely serious.

These statistics were used to determine the strength and significance of correlation between the quality framework as a whole and program performance outcomes. This correlation served as a measure of the criterion-related validity of the framework. Additionally, data from

the feedback form were used to supplement the correlation due to the small sample size and provide evidence of face validity.

Question two. To answer the second research question of this study, “Is the evaluation instrument a useful tool for local educators to assess the quality of their CTE programs?” quantitative and qualitative data from the feedback questionnaire accompanying the evaluation instrument were examined. Frequency distributions and simple descriptive statistics such as means and modes were used to analyze quantitative data about the evaluation experience and other potential use. Content analysis techniques were used to identify trends in open-response questions that yielded qualitative data.

Phase Two

Sample. The sample for the second phase of the study consisted of five individuals sampled purposefully for one-on-one interviews with the researcher from those that submitted programs during Phase One. Criteria for selection included geographic diversity, CTE program diversity, and educational setting diversity, including programs in high schools, area CTE centers and postsecondary institutions. These individuals were also selected to represent a range of scores on the evaluation instrument in Phase One, to ensure diversity of perspectives. Due to the difficulty that many individuals had completing the entire study, and the time lapse since the first submissions were received, participants for Phase Two were also selected based on those who had shown continued interest and enthusiasm for the project and seemed likely to participate and have valuable input. For example, one participant was selected because she went to great lengths to obtain the necessary Perkins data to allow her program to be included. Finally, since the validity data obtained in Phase One were not as robust as originally desired, individuals who submitted multiple programs to the study were selected for Phase Two so that they could be

asked about comparative results. All individuals originally asked to participate completed their interviews in Phase Two.

Instrumentation. An interview protocol was developed to gather perspectives on the utility and scale-up potential of the evaluation instrument. The interview questions for this purpose were developed after completion of Phase One of the research, and were based on inferences and trends identified during initial data analysis. Questions were tailored to each individual based on their results from Phase One. Below are examples of questions that were used for related inquiry:

- Do you think the results your program received were valid and accurate?
- Do you feel that the evaluation instrument completion was a useful exercise? If so, what insights did you gain about your program?
- Are there additional tools and resources you would have liked to receive along with the evaluation instrument to make it more useful?
- Do you think that other CTE programs at your institution would find the tool helpful?

Data collection procedures. Data were collected under the second phase of the project through a set of 30 minute in-depth telephone interviews with the participants selected for the sample. Interviews were used due to their adaptability, which is suggested by Gall, Gall, and Borg (2007) to be their biggest strength. Each interview was tailored to the specific participant and was responsive to the information they provided, allowing for additional depths of information to be obtained. Telephone interviews were used rather than in-person interviews due to the geographic diversity of the sample and high costs that would be associated with travel.

Interviews occurred during the summer of 2018, after the majority of Phase One work and initial data analysis was complete. Participants were provided with copies of the original

information they submitted as a refresher due to the lapse of time since some of them had completed the Phase One instruments. The interview protocol developed for this purpose was used to gather the participants' thoughts on the evaluation rubric completion process, its utility and implications for wide spread use. Participants were provided with a conference call number to participate in the interview. At the beginning of the interview, participants were reminded about ACTE's overall High-quality CTE Initiative. Then, participants were debriefed about the purpose of the study and interview. Next, assurances about confidentiality and voluntary participation were reiterated, and in turn, permission to record the interview was sought. After consent and permission to record was granted, then conference call service recording was started, and participants were asked to provide verbal consent again on the recording. After the completion of the interview, recordings were downloaded and transcripts were created for future analysis.

Data analysis. Building on Phase One of the study, interview data were used to supplement data on both research questions about instrument validity and utility. The interview transcripts were analyzed using a thematic content analysis approach, which Braun and Clarke (2006) define as “a method for identifying, analysing, and reporting patterns (themes) within data” (p. 6). They suggest six steps for thematic analysis:

1. Becoming familiar with the data: This step involves repeated, active reading to immerse the researcher in the depth and breadth of the content.
2. Generating initial codes: In this step, the data begins to be organized with initial codes representing interesting features of the data.
3. Searching for themes: The third step of the process involves refocusing the work more broadly and organizing the individual codes by themes.

4. Reviewing themes: The fourth step of the process involves reviewing coded data and then the entire dataset again in relationship to the draft themes to ensure proper fit, with edits made to the themes as needed.
5. Defining and naming themes: The fifth step in the process involves developing a narrative analysis of each theme from the data, as well as formal “titles.”
6. Producing the report: The final step in the process involves developing the final analysis and reporting of the data from this phase of the process.

Finally, and in keeping with the mixed-method design, inferences from the totality of data collected and analyzed in both Phase One and Phase Two of the project was reviewed and examined to draw final conclusions. In this regard, the mixed-method approach provided greater depth in response to research questions than either method alone could accomplish.

CHAPTER 4: RESULTS

The results of both Phase One and Phase Two of this study are included in this chapter. First, an overview of participation in the evaluation instrument and survey are presented. Then results from Phase One are reported, with data from the evaluation instrument and survey related to the research question on the validity of the evaluation instrument's results. Results related to the question of utility of the instrument are also addressed. Finally, results related to Phase Two interviews are presented to add practitioners' insight about the validity of the evaluation instrument's results and its utility.

Phase One

Description of sample. Data collection began in December 2017 and continued through August 2018. Despite significant outreach encompassing emails to over 20,000 members of ACTE, presentations at numerous conferences, and multiple targeted follow-up attempts, the availability of the Perkins data necessary for the study proved to be a significant barrier to participation, even for those who were very interested in the evaluation instrument itself.

For Phase One of the study, a total of 56 submissions of the evaluation instrument were received, the first step of the process. However, of these, only 39 included the actual program evaluation data (the remainder only included contact info). At least eight of the incomplete evaluations appeared to be mistakes, as those individuals returned later and submitted a complete form.

The unit of analysis for this phase of the study was the individual program, so very little information was collected about the program representative who completed the instrument—only

theirs names and email addresses for potential follow-up. For this reason, the descriptive information included here focuses on the programs themselves. The 39 completed evaluation instruments reflected a diversity of CTE program delivery systems. Submissions were received from comprehensive high schools, magnet high schools, school districts, community colleges, CTE centers, and colleges. A total of 21 states/territories were represented in all five regions of the country (as defined by ACTE). Ten evaluations were received from postsecondary programs and 24 from secondary programs, with the remainder unable to be identified by the information provided due to being from institutions that serve both secondary and postsecondary students.

Twelve of 16 Career Clusters were represented, with Health Sciences representing the largest share of programs. The number of participating programs from the Health Sciences Cluster was significantly more than other Career Clusters, and could have had an impact on the results. On the one hand, programs in this area are known to have close connections with industry and clear standards, which could lead to higher scores on the evaluation instrument. However, there is also a benefit to examining the impact of nuances in program quality across a range of more standard programs. Table 4.1 shows the number of programs participating per Career Cluster.

It is important to note that the 39 programs represented in this study are a very, very small percentage of the total population of CTE programs around the country. There is no national data reflecting this total number, but CTE programs serve 94% of all of the country's high school students, with approximately 20% of those participating in a sequence of three or more courses (U.S. Department of Education, National Center for Education Statistics, 2009). Approximately 12 million postsecondary students participate as well (ACTE, n.d.). Most

schools that offer CTE offer multiple, individual programs, and community and technical colleges or area CTE centers might individually offer dozens of distinct programs.

Table 4.1

Program Evaluation Instrument Participants by Career Cluster

Cluster	Participants
Agriculture, Food & Natural Resources	1
Architecture & Construction	4
Arts, A/V Technology & Communications	1
Business Management & Administration	3
Education & Training	2
Finance	0
Government & Public Administration	0
Health Science	10
Hospitality & Tourism	3
Human Services	2
Information Technology	5
Law, Public Safety, Corrections & Security	2
Manufacturing	0
Marketing	0
Science, Technology, Engineering & Mathematics	1
Transportation, Distribution & Logistics	5

Only 34 responses were received from the second instrument, the evaluation feedback form. Of these, only 29 included enough information to link the two sets of results, and only 23 were able to provide usable Perkins data (8 postsecondary and 15 secondary). Obtaining program level Perkins data proved to be extremely difficult for some participants who attempted to participate, for a variety of reasons, despite specific directions that it was required for participation. Some of the programs did not have data because they were new; were in an institution that did not receive Perkins funds; had data delay issues; were too small for their state to make data available; or could not find the person in their school, district or state who could provide the disaggregated numbers.

Overall, while the sample was not as large as originally desired, with only 23 complete data sets, it reflected a breadth of program types, geographic locations, and institutions. While the low number of programs that were able to submit Perkins data impacted the results on the evaluation instrument's validity, the fact that many of these programs still participated in the study helped to address questions of utility and increased the relevance of that information and its applicability across the CTE system.

Evaluation instrument results. Completion of the program evaluation instrument was the first step of the research process, and results on the instrument were used to address the question, "Does the program evaluation instrument provide valid evidence of program quality?" On the program evaluation instrument, participants in the study were asked to rank their programs on 102 criteria grouped into 12 elements on a scale of 0 to 3. On the scale, "0" represented that the criteria was "not at all achieved," "1" represented that the criteria was "minimally achieved," "2" represented that the element was "moderately achieved" and "3" represented that the criteria was "substantially achieved." Additional descriptors were provided to explain these four categories labels and can be found in Appendix C.

The number of criteria in each element varied and is noted in Table 4.2, which contains element-level score details. The possible total score range was 0 to 306; while the actual score range was 145 to 300. For the purposes of calculating the total score for each program, there were minimal missing data points (12 out of the 3,978 total possible), and those were treated as "0." The mean of all 39 total scores was 237.79, with a standard deviation of 42.71, and the median score was 242. An internal consistency measure, Cronbach's alpha, was also calculated using only the 35 complete data sets (containing responses on all 102 items), and was found to be .97. When comparing the two groups of submissions—the 23 programs that were able to

provide Perkins data and the 16 that did not, it was found that they were very similar practically. The range for the group that did provide Perkins data was 145 to 298, the mean score was 236.13, and the standard deviation was 40.52. For the 16 programs that were not able to provide Perkins data, the range was 153 to 300, the mean score was 240.19, and the standard deviation was 46.93.

Descriptive statistics on scores for criteria within each of the individual 12 elements were also calculated, and mean element-level scores ranged from 1.93 to 2.74 (with a possible range of 0 to 3). Total scores for each element were not calculated due to the different numbers of criteria within each element, which made comparisons of total element-level scores meaningless. On average, the highest scoring element was “Facilities and Equipment,” and the lowest was “Data and Program Improvement,” which echoes the challenges participants had obtaining the necessary data for participation in the study. The mean scores for criteria within each element, in order from lowest-scoring to highest-scoring, are reported in Table 4.2, along with other descriptive information. In addition, Cronbach’s alpha was used to determine the internal consistency of each element’s specific criteria scores. The Cronbach’s alphas ranged from .74 to .94. Those values for each element are also included in Table 4.2.

Validity of evaluation instrument. As mentioned above, only 23 programs (eight postsecondary and 15 secondary) were able to provide the Perkins performance data necessary to be included in the correlation between the total score on the evaluation and program outcomes. Even within the 23, some programs were missing some data elements. Of the three outcome indicators chosen, one was a complete set, and the other two were missing three scores each (although not the same scores).

Table 4.2

Program Evaluation Instrument Results: Information on Scores for Criteria Within Each

Element

Element	# of Criteria	# of Total Scores	<i>M</i>	<i>SD</i>	α
Data and Program Improvement	9	351	1.93	.65	.91
Career Development	8	312	2.01	.63	.90
Business and Community Partnerships	14	546	2.17	.71	.94
Career and Technical Student Organizations	8	303 ¹	2.21	.99	.97
Work-based Learning	12	468	2.25	.70	.94
Sequencing and Articulation	7	273	2.31	.63	.85
Student Assessment	8	311 ¹	2.43	.54	.89
Standards-aligned and Integrated Curriculum	7	273	2.54	.46	.74
Prepared and Effective Program Staff	7	273	2.56	.40	.75
Access and Equity	7	272 ¹	2.62	.37	.79
Engaging Instruction	7	272 ¹	2.63	.39	.79
Facilities and Equipment	8	312	2.74	.34	.80

¹These elements were missing at least one score across criteria/responses. For the purposes of calculating the mean and standard deviation, missing scores were treated as missing individually (and all other criteria scores across all responses were used). For the purposes of calculating Cronbach’s alpha, all scores had to be present across an entire element for responses from that participant to be included.

The indicators used in the correlation included: technical skill attainment, completion, and placement. Each of these indicators represents data reported under the Carl D. Perkins Career and Technical Education Act by local Perkins funding recipients to their respective states, and then to the federal government in aggregate form. Each state defines the specific definition of the numerator and denominator that is used to create a percentage between 0 and 100 that becomes the reported indicator. Generally, the denominator is the total population of CTE concentrators that left the program in a given year (either by graduating or dropping out), and the numerator is the percentage of those students who achieved the specified goal. States also define the specific measurement approach. For the technical skill attainment indicator, the measure varies, but may include achievement of an industry certification or performance on a statewide

technical skill assessment. Completion represents either high school or postsecondary credential completion, and placement represents further placement in education or employment. The data for each indicator were self-reported by participants on the feedback questionnaire.

The Pearson product-moment correlation was calculated for each of the indicators above paired with the participant’s overall score on the evaluation instrument. Correlations were positive in all areas, and the results are reported in Table 4.3.

Table 4.3

Correlation Results

Indicator Paired with Overall Score	<i>n</i>	<i>r</i>	<i>p</i>	<i>r</i> ²
Technical Skill Attainment	20	.41	.07*	.17
Completion	23	.46	.03*	.21
Placement	20	.15	.54	.02

**p* < .10

The first correlation was between the overall score on the evaluation instrument and the percentage of students achieving technical skills. These two variables were moderately correlated, with a significant positive correlation between the results on the evaluation instrument and technical skill attainment, $r(18) = .41, p < .10$. Based on the coefficient of determination, almost 17% of the variation in technical skill attainment scores was accounted for by the variation of scores on the evaluation instrument. As the overall score on the evaluation instrument increases, a moderate increase in the number of students reaching the technical skill attainment benchmark could be expected.

The second correlation was between the overall score on the evaluation instrument and student completion (earning the required credential). These two variables were also moderately

correlated, with a significant positive correlation between the results on the evaluation instrument and student completion, $r(21) = .46, p < .10$. Based on the coefficient of determination, over 21% of the variation in completion was accounted for by the variation of scores on the evaluation instrument. As the overall score on the evaluation instrument increases, a moderate increase in the student program completion could be expected.

The last correlation was between the overall score on the evaluation instrument and post-program placement results. These two variables had only a small correlation, which was not significant, $r(18) = .15, p > .10$, with only 2% of the variation in placement results accounted for by the variation of scores on the evaluation instrument.

Based on the analysis of correlation results, despite the small sample size, there is some evidence of the instrument's validity in relationship to program outcomes as reported under Perkins, particularly for technical skill attainment and student completion.

Correlations were also calculated between data under each of the 12 elements and the three measures. These results can be found in Appendix G. There was no strong pattern within the results, other than the fact that no elements had large correlations with "placement," which had the lowest overall correlation as well. More elements were negatively correlated with placement as well. The elements that had the weakest relationship across all three outcomes indicators were "Prepared and Effective Program Staff," "Engaging Instruction," and "Career and Technical Student Organizations."

In addition to the correlations used to address the instrument's criterion validity, several questions were asked on the feedback form to address face validity of the instrument and supplement the statistical analysis. In response to a question on how well responses and scores on the Evaluation Instrument reflected the overall quality of the program being evaluated, the

majority of respondents said “very well” (26.47%, $n=9$) or “well” (44.12%, $n=15$). No participants reported that the instrument reflected their program “very poorly,” and only one said the instrument “poorly” reflected the program’s quality. Overall, 92.14% of all respondents also agreed that the instrument covered the key elements of their program.

Utility of the evaluation instrument. All 34 responses that were received on the feedback questionnaire were analyzed and included in the results, even if these responses did not include program information or Perkins data. Responses to questions regarding the utility of the instrument were overall quite positive.

Ease of use. Respondents were asked to assess the evaluation instrument on several factors related to its ease of use. A five-point Likert scale, ranging from “very poor” to “very good” was used to collect results. Overall, each factor was rated on average at least “good,” with most being between “good” and “very good.” The instructions as a whole was the factor rated the highest, with 58.82% of respondents rating them “very good,” followed by “time to complete,” with 45.45% of respondents rating this factor “very good.” However, time to completion was the aspect that drew the most additional comments, with seven respondents mentioning either the length of the survey specifically or offering related comments, such as the desire for a progress bar or the suggestion that some of the indicators could be deleted due to redundancy. Full results from these ease of use ratings are presented in Table 4.4.

Respondents were also asked how easy or difficult it was to find the information to complete the evaluation, and the majority responded positively, with 38.24% ($n=13$) responding “very easy” and 41.18% ($n=14$) responding “easy.” Some comments suggested that teachers alone might have trouble with some of the questions though. For example, one respondent (an administrator) said, “My teachers didn't understand some of it, but with my help it was easy

Table 4.4

Evaluation Instrument Ease of Use Ratings

Factor	<i>n</i>	Very Poor 1	Poor 2	Fair 3	Good 4	Very Good 5	<i>M</i>	<i>SD</i>
Instructions for Use	34	0% 0	0% 0	11.76% 4	29.41% 10	58.82% 20	4.47	0.71
Number of Criteria	34	0% 0	2.94% 1	17.65% 6	44.12% 15	35.29% 12	4.12	0.81
Rating Scale Provided	33	0% 0	3.03% 1	24.24% 8	42.42% 14	30.30% 10	4.00	0.83
Time to Complete	33	0% 0	0% 0	21.21% 7	33.33% 11	45.45% 15	4.24	0.79

enough.” However, 15 of 33 responses to a separate question suggested that participants did not reach out to anyone else to help, and several more seemed only to need assistance to retrieve the Perkins data (which were not actually part of the evaluation instrument). Others mentioned consulting with their CTE director/coordinators, program instructors, or various co-workers.

On a wrap-up question on ease of use, 41.18% ($n=14$) of respondents rated the instrument “very easy” to use overall, another 44.12% ($n=15$) rated it “easy” to use overall, and 14.71% ($n=5$) gave it a neutral rating on overall ease of use. No participants rated the instrument difficult or very difficult to use.

Usefulness of instrument. The next set of questions asked respondents about the usefulness of the instrument. First, respondents were asked whether they gained new insights into improving their program from completing the instrument. 82.35% ($n=28$) responded “yes” to this question, and 17.65% responded “no.” Comments suggested that participants picked up specific ideas for how to improve their programs, and engaged in conversations they might not

otherwise have had about program elements. The only comment from a participant who responded “no” expressed that they didn’t think they learned much because their program was so high quality, but that programs with more room for improvement would find the tool useful. Table 4.5 contains the representative sample of comments related to what respondents learned, and illustrates how the instrument changed their perspective on several areas.

Table 4.5

Insights into Improving Programs

Comments Received	
Need to develop more work based learning experiences for students (High school teacher)	I really thought about how I could better utilize my program advisory board (High school teacher)
Confirmed our strength in resources and materials ... identified our need for growth in the area of certifications (High school teacher)	Need a stronger career education plan outside of our teachers in the POS (High school teacher)
Opportunity to discuss with program faculty (Community college administrator)	Realizing that my advisory committee has power in giving my program input (Community college instructor/program coordinator)
Opportunity to discuss with program coordinators (Community college administrator)	I believe that based on our choice of program (well established, overall very successful with a good track record) the process was straightforward and did not reveal any insights, but if a program that was less successful or struggling was analyzed it could provide feedback as to the areas on which to focus (Tech center administrator)
Using the quality standards for our annual evaluation (School district CTE director)	
This assisted us in identifying the weaknesses we sensed (Tech center administrator)	
Our department has a lot to work on to ensure that we deliver high-quality CTE program (Community college department chair)	Need to work with faculty on definitions and their perceptions of quality indicators (Tech center administrator)
I think there are a lot of good expectations, but some of the items are beyond control of our institution or in some states, state laws (Technical college administrator)	This instrument highlighted the need to share data with staff, students, and parents/guardians. (Tech center administrator)
How the Advisory needs to have more input, plus the data needs to be used for improving the POS (High school teacher)	Mainly, to keep promoting and working on business/industry partners. Industry for the Teaching/Learning program is education and education-related industry (High school teacher)

Overall, respondents rated the tool useful to evaluate and improve program quality, with over half, or 55.88% ($n=19$) reporting the instrument was “very useful.” An additional 38.24% ($n=13$) rated it “moderately useful,” and only one respondent each (2.94%) rated it “slightly useful” or “not at all useful.” In a comment about the instrument’s utility, one respondent summed up why they rated the tool so highly:

We have been using a state-wide evaluation tool but it doesn't break down the indicators as completely as this tool does. We have also generated our own survey tool. I would have all CTE teachers use this survey for their programs along with site, district leadership and then compare. From that develop a plan to address areas for growth and annual goals. (CTE director for a secondary school district)

In general, the quote above suggested that the comprehensive nature of the instrument was appreciated by practitioners, who saw the potential value for use as an evaluation or survey tool to learn how programs are doing.

Implications for future use. The next question dealt with how the evaluation instrument should be used in the future. Respondents were allowed to select as many of the items they felt appropriate from a list of eight possible uses. Teacher professional development was the response selected most often, by 27 (81.82%) respondents. The next most common responses were highlighting best practices and technical assistance to local programs, each with 23 (69.70%) responses. Allocating funding was the least common response, with only eight (24.24%) respondents selecting this as an option for how the tool should be used in the future. Full results are in Figure 4.1.

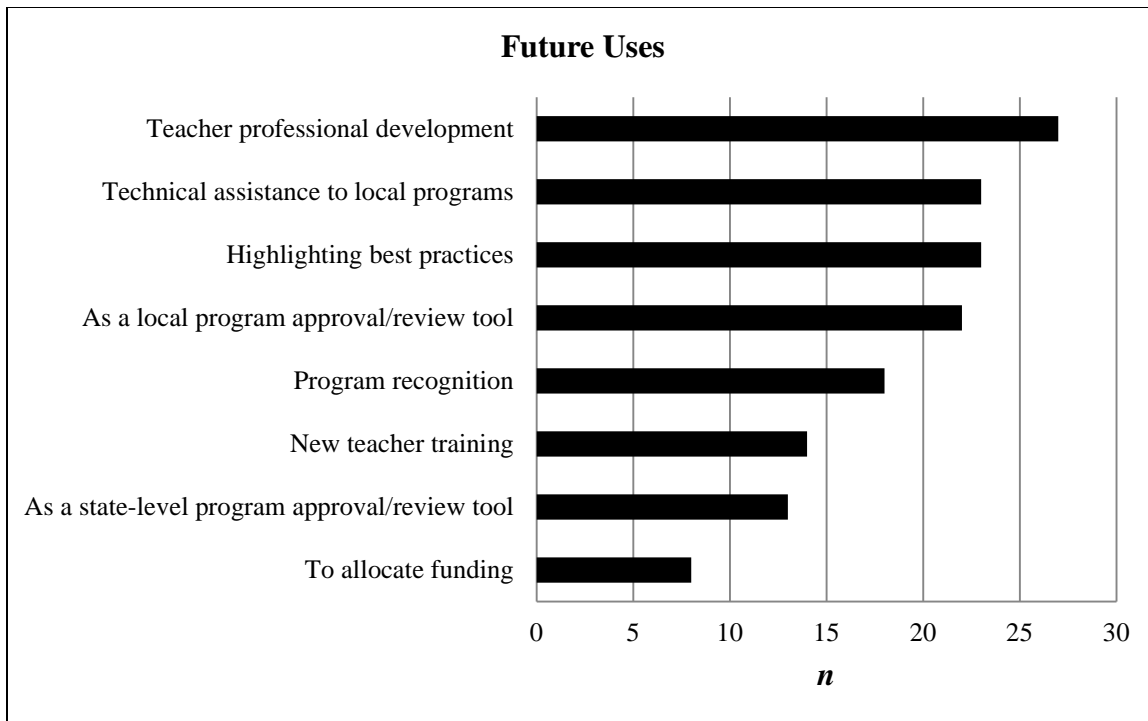


Figure 4.1 Suggestions for Future Use

Practicality across CTE program models. As stated above, almost all participants agreed that the instrument covered the key elements of the program, regardless of the type of CTE program they offered. Digging deeper into the applicability of the instrument across program models, only about one-third, or 32.35% ($n=11$), suggested there were elements that did not apply to their program, and about that same number, 33.33% ($n=11$) suggested additional elements that would be helpful to include. Related to elements that did not apply, some respondents highlighted things they do not do in their program, like career and technical student organizations, while others highlighted areas that they felt were out of their control, such as vertical program alignment or the work of school counselors. A few concerns were also raised related to specific high school-focused language. For example, one respondent from an area CTE center postsecondary program said that “communication with parents for adult programs is

not relevant.” Suggestions for new elements to add varied, but there were two mentions of adding a question on administrator support and other suggestions to include space for elaboration or explanation.

One final question gave respondents a chance to offer any more suggestions for changes to allow the evaluation instrument to be used by more programs. Only 16 respondents provided a response to this open-ended question. Of those, six said they had no suggestions for improvement. The remaining 10 suggestions primarily covered ideas already raised around three areas of potential improvement:

- The online interface needs to be improved—for example, there was a suggestion for a “save and finish later” option. Within the online interface used for the pilot, the instrument did not allow for saving temporarily and returning to finish at another time.
- Including an opportunity for respondents to provide more detail would be helpful, such as through open-ended responses opportunities, or allowing the attachment of evidence. A suggestion was also made for a suggested list of evidence for each indicator.
- Some criteria may need to be altered for different education levels (middle school, secondary, postsecondary). For example, participants referenced mentions of involving parents as inappropriate for postsecondary students.

Phase Two

Phase Two of this study was designed to build on the findings and conclusions of Phase One, based on the sequential, mixed model design as described by Tashakkori and Teddlie (2003). It was used to gain additional insights into both research questions through in-depth phone interviews, yielding additional information on the validity of the instrument’s results and its utility.

Description of sample. Five individuals who had submitted data on programs under Phase One were selected to participate in Phase Two of the study, and represented a range of CTE institution types. One was from a community college, two from secondary school districts working with programs at comprehensive high schools, and two from area CTE centers. All participants were administrators at the school or district level, selected purposely for their enthusiasm for the project and potential for having significant input. Two of the individuals had submitted two different programs each under Phase One of the study, providing them with additional experience with the tool. The states represented included Arizona, Colorado, Illinois, Indiana, and a small state that will not be named to ensure the privacy of the participating individual.

Emergent themes. The interviews focused on building on the results of the first phase of the project in order to dig deeper into the questions of the evaluation instrument's validity and utility. Since results on the first phase of the project were very positive for the instrument's ease of use and usefulness, the interviews attempted to go beyond these basic questions and gather respondents thoughts not just on their experiences, but on future utility of the instrument. In addition, since the validity analysis was weakened by a lack of participants, questions about the instrument's perceived validity were also included. After analysis of the interview transcripts using an inductive approach to thematic content analysis, the following themes emerged and are elaborated on below:

- The instrument is seen as a valid measure of program quality.
- Administrators and educators may have different perspectives.
- This instrument fills a gap in available quality metrics.

- The use of this instrument lends credibility to participating programs.
- The instrument is an effective program improvement tool across the CTE spectrum.

The instrument is seen as a valid measure of program quality. To supplement data from the statistical analysis in Phase One, participants were asked about the scores their programs received on the evaluation instrument and whether they perceived those scores as valid. Across all five interviews, scores matched participants' overall thoughts about their programs and suggested the instrument represented a valid measure of program quality. For example, two participants who had submitted data on two programs were asked in particular about their respective program outcomes and agreed that the program rated higher was actually higher quality. One elaborated as follows:

It makes absolute complete sense. The program has been around longer, we have more partners in the community supporting the program, we have a person that has a portion of her FTE designated to coordination of that program as opposed to the other one is the teacher alone, you know, doing the coordination of the partnerships with the local community. (Area CTE center administrator, Illinois)

The other participant with two programs also agreed the higher-scoring program was higher-quality. "I'm not surprised by the scores," the community college dean from the small state said, and elaborated on the higher-scoring program's structured curriculum, high-quality faculty, and national industry accreditation. This perspective was further reinforced by another participant who submitted data on a very high-performing, nationally award-winning program. This program's score was one of the highest in the pilot, but in breaking down element subscores, the participant agreed those were areas needing improvement. The district CTE director from Arizona said, "Even though it's a quality program, there's always room for

improvement, especially in career development...I thought that [the lowest subscores] was pretty accurate.” Another participant, a district CTE director from Colorado, originally thought her program would score high, but as she dug in, realized areas of work were needed. “It did surprise me how low I rated it, but I think it's accurate,” she said.

The area CTE center administrator from Indiana addressed the elements themselves and their representativeness. He mentioned that after working with a team to evaluate their program, that other staff “feel like those 12 points define what it means to be a quality program effectively.” He went on—“It's not just me commenting. There are others as well!”

Administrators and educators may have different perspectives. While all of the interviews were conducted with administrators in order to ensure the best participation, it became clear from the conversations that teachers and administrators might have different perspectives on program quality due to their unique roles and responsibilities—with educators having more direct classroom experience, but administrators offering a broader perspective.

First, it was noted by participants that educators might have a natural bias toward believing their programs were doing better than they actually were. One participant, the area CTE center administrator from Indiana, said that “teachers may be susceptible to a little bias in their reporting” if using the instrument as a self-reflection tool, and the community college dean from the small state said, “If I had given this to a program coordinator, maybe some of the answers might be different.”

In addition to an inherent bias, several participants mentioned other reasons why scores from teachers and administrators might be different. One reason was not enough consistency in what is expected of or articulated for educators related to quality. “I think for sure our teachers do [have an overinflated sense of program quality] and I think it's just because there's not enough

consistency and this is what it looks like, you know, we just throw out general terms....” said the district CTE director from Colorado. Another related reason was a lack of comparative information about program quality. The district CTE director from Colorado went on to provide another example where a teacher might think they have a quality work-based learning program because 17 students had done job shadowing experiences, but another teacher might have placed that many students in full paid internships. Without specific, comparative information, the first teacher would not know there were gaps in their program’s quality.

Classroom-based educators also may not have as much exposure to the language used in the rubric or all the information needed to complete it. Participants who consulted with their program’s teachers or faculty raised the issue that there were criteria that needed more explanation or information. “Had I not been there to further explain what the question was asking, she wouldn’t have known what to say,” the area CTE center administrator from Illinois said, and went on to suggest a glossary of terms or some linked framing documents to accompany the instrument.

From the opposite perspective, interview participants suggested that administrators might not always have the information necessary to complete detailed questions on the evaluation form. Even those interviewed mentioned looking up data, consulting with educators, and the importance of actually visiting programs and seeing them in action. External administrators “may have little familiarity with a number of these questions,” the area CTE center administrator from Indiana explained. The area CTE center administrator from Illinois added, “I feel like I have one view, but there’s always more than one story to tell,” pointing to the importance of including multiple perspectives in any program evaluation. This also led to suggestions for

external validation of results if the evaluation instrument was going to be used in any high stakes way.

The instrument fills a gap in available quality metrics. One of the primary purposes of this research was to determine whether the evaluation instrument was useful for program improvement. In digging into that question, it became evident that current quality metrics do not meet the needs of program in identifying strengths or areas for improvement, and that this program evaluation tool can help fill that gap. Three primary points were raised related to this theme.

The first issue identified was that the tools many local administrators have to use to evaluate programs or teachers don't address any CTE specific information. The area CTE center administrator from Illinois mentioned that their teacher evaluation tool, the Danielson model, evaluates CTE teachers, but "has nothing to do with the actual program components." The district CTE director from Colorado added that principals are often asking for help because "the teacher rubric for CTE programs is the same as all the other programs, which in a way, you know, is fair—that's equitable, but in a way it doesn't really match what we do in our classrooms."

The next issue raised related to this theme was that in local control states, benchmarks for program quality are lacking. The district CTE director from Colorado explained that while her state office provided resources and support, it was up to the district to "help their schools and their program know what, what's happening with CTE and what is expected" and that having external guidance was very helpful to determining what should be expected. The area CTE center administrator from Illinois mentioned that rather than relying on her own knowledge, the framework served as "a good objective tool."

Finally, participants raised the idea that relying on Perkins performance indicators or other outcome measures was not enough to really improve programs. “We have the Perkins performance indicators to look at, but yet, you know, we also need to interpret what the numbers and the data mean,” said the community college dean from the small state who went on to elaborate that this evaluation tool would be very helpful to truly evaluate the health of programs. The district CTE director from Arizona summed up the need for a broader, more in-depth focus on quality rather than just compliance, “We knew that we had to do an annual evaluation for Perkins compliance, but just evaluating on performance measures data did not drive our ships forward.”

The use of this instrument lends credibility to participating programs. Participants repeatedly highlighted the additional credibility that the use of the program evaluation tool, with its underlying research and national endorsement, gave them or their programs with a variety of audiences. This respondent summed up this theme:

So using this I think helped me to see what other people around the nation are doing and how they're framing a high-quality program. And seeing, you know, all the research that you've done—it's not like you just took a survey and you said, “Hey guys, what do you think is a good program?” There was a ton of research and a ton of work that went into this and so knowing that I felt really comfortable using it as that tool to say, “Okay, let's look a little bit at where we can grow our programs, how we can do better.” (District CTE director, Colorado)

Educators were one stakeholder group where buy-in was highlighted, with the area CTE center administrator from Illinois suggesting that teachers were much more receptive because it was a national organization, “...and so instead of it just coming from me, I selfishly was

thinking, hey, here is a professional organization who's saying here are quality components.” The district CTE director from Colorado also expressed that she thought her teachers would benefit from the national perspective. “I’m really excited to use it more to help my teachers see what nationally is considered high quality CTE,” she said.

Additionally, participants discussed the value of participating in a national evaluation, mentioning that they often seek out program certifications that require them to prove competency. An ACTE endorsement was mentioned by participants as a possible next step in this regard, although participants seemed to value the results alone as well. Local school boards, other administrators, policymakers and employers were all mentioned as possible audiences to share results with. The area CTE center administrator from Indiana suggested a program endorsement might be useful with these groups, elaborating, “I think there's significant value in being able to go to a local school board or go to employers and say, ‘Hey, you know, this is a quality program as evidenced by this endorsement.’” The community college dean from the small state echoed those thoughts, “I think that would certainly help in [small state] and I imagine nationally too, you know administrators and the legislators, industry...They're all looking if the program is, you know, healthy,” and suggested that the evaluation instrument would be a valuable part of that conversation. Other participants suggested using the credibility of the national document to inform strategic plans, share with other colleagues, or determine resource allocation.

The instrument is an effective program improvement tool across the CTE spectrum.

Across the board, participants agreed the evaluation instrument was an effective program improvement tool, and met the needs of a variety of programs—from high school to community

college, and more traditional classroom-based programs to apprenticeships. One participant shared his hopes for wide-scale use:

So this kind of instrument certainly is great for all of our seven community colleges here in [small state] to evaluate their programs and we have monthly, we meaning the career and technical education deans, have monthly meetings that it'd be a good opportunity for us to have this as a standing agenda item to talk about our programs, how we evaluate the programs and the, as far as the health, you know, of our programs and how we can improve our programs throughout the [small state] system of community colleges.

(Community college dean, Small state)

Participants went on to describe the instrument as “a valuable process,” “an enhancement to what we already do,” and useful for “professional growth.” For example, the area CTE center administrator from Indiana identified the tool’s results as “a catalyst” for further discussion and change, and the community college dean from the small state “appreciated the opportunity to compare” results across programs. Having measurement criteria for all of the diverse elements was also noted as an important aspect. “What’s measured gets done. So if we have a measurement criteria on it, then we have a target that we can work on to improve,” said the district CTE director from Arizona.

Participants cited a key benefit of the tool as drilling down into programs and identifying those key areas that needed additional focus or improvement. The area CTE center administrator from Indiana explained, “I think most of our programs are pretty strong in each of these areas, but some may be lacking miserably in one or two key areas...I think it's more effectively used to identify, you know, the one, two or three priority areas that we should be focusing on for the next couple years to get that program up where it needs to be.” In turn, the community college dean

from the small state suggested this ability to identify key areas would be particularly helpful to administrators in determining priorities, “One of the strengths I think for the evaluation tool is...it affords myself an opportunity to go back to our division chair and program coordinators to see do we need to revisit our curriculum or do we need to revisit our professional development opportunities.” Participants also suggested that low scoring criteria could be connected to resources to drive program improvement. The district CTE director from Colorado suggested ACTE could say, “Your program scored low in this area. Here are some resources you can use to help grow the program or to help that teacher.”

CHAPTER 5: DISCUSSION

This chapter is organized into five sections. First, an overview of the study is provided to set the context for discussion. This is followed by a summary of the results of the research. Then, the findings are discussed in relationship to prior research and the conceptual framework, and conclusions about the evaluation instrument are presented. Limitations of the study are presented next, and finally, implications for future practice and research are discussed.

Study Overview

The purpose of this study was to build on the research already completed by the Association for Career and Technical Education (ACTE) in the development of the ACTE high-quality Framework and accompanying evaluation instrument. The goal was to examine the instrument's validity and utility when used independently by local educators in a pilot test. The specific research questions were:

1. Does the program evaluation instrument provide valid evidence of program quality?
2. Is the evaluation instrument a useful tool for local educators to assess the quality of their CTE programs?

In Phase One of the research, participants used the instrument to evaluate one of their local programs, then were asked to complete a feedback form and provide Perkins performance data on that program. Results of correlations between the scores on the evaluation instrument and individual performance indicators were analyzed to examine criterion validity—whether the included elements are appropriate measures of program quality, as related to already established Perkins program outcome measures. In addition, results on the feedback form were analyzed to

examine face validity, as well as the instrument's utility, including ease of use, practicality, and usefulness of results to local practitioners. In Phase Two of the study, results of Phase One were used to inform in-depth interviews with a sample of five participants. These interviews provided further insight on the validity and utility questions of Phase One, including the resources, barriers and possibilities for scaling up the tool to widespread use.

Summary of Results

Evaluation instrument results. The actual score range on the evaluation instrument was 145 to 300 (on a possible scale of 0 to 306). The mean of all 39 total scores was 237.79, with a standard deviation of 42.71, and the median score was 242. Cronbach's alpha was calculated using only the 35 complete data sets (containing responses on all 102 items), and was found to be .97. Descriptive statistics on scores for criteria within each of the individual 12 elements were also calculated, and mean element-level scores ranged from 1.93 to 2.74 (with a possible range of 0 to 3). On average, the highest scoring element was "Facilities and Equipment," and the lowest was "Data and Program Improvement." Cronbach's alpha was used to determine the internal consistency of each element's specific criteria scores. The Cronbach's alphas ranged from .74 to .94.

Research question one. *Does the program evaluation instrument provide valid evidence of program quality?* To determine the criterion validity of the instrument, a program input measure—in this study, the participant's score on the program evaluation instrument—were correlated with program outcome measures—which in this study were Perkins performance measures. The overall evaluation instrument score was correlated individually with three Perkins performance measures: technical skill attainment, completion and placement.

Due to a small sample size and low power, a larger alpha of $p < .10$ was used to test for significance and reduce the Type II error chance (Grimm, 1993). Results for the correlation with technical skill attainment showed a statistically significant, $r(18) = .41, p < .10$. Based on the coefficient of determination, almost 17% of the variation in technical skill attainment scores was accounted for by the variation of scores on the evaluation instrument. There was also significant positive correlation between the results on the evaluation instrument and student completion, $r(21) = .46, p < .05$. Based on the coefficient of determination, over 21% of the variation in completion was accounted for by the variation of scores on the evaluation instrument. Finally, the last correlation was between the overall score on the evaluation instrument and post-program placement results. These two variables had only a small correlation, which was not significant, $r(18) = .15, p > .10$. Only 2% of the variation in placement results was accounted for by the variation of scores on the evaluation instrument. The results suggested that a higher overall score on the instrument may be an indicator of program quality in terms of skills attainment and student completion. However, the score on the instrument did not appear to be associated with program-placement results as measured by Perkins data.

The results on criterion-related validity were supplemented by data on face validity from the feedback questionnaire. The majority (70.59%) of respondents said the instrument reflected the quality of their program “very well” (26.47%, $n=9$) or “well” (44.12%, $n=15$). A full 92.14% of all respondents also agreed that the instrument covered the key elements of their program.

Finally, face validity was also addressed in the Phase Two interviews, where one of the themes that emerged from the interviews was that participants saw the instrument as a valid measure of program quality. Participants were asked about the scores their programs received on the evaluation instrument and whether they perceived those scores as valid. Across all five

interviews, scores matched participants' overall thoughts about their programs. Individual participants that had submitted data on multiple programs also agreed that the higher scoring program was indeed higher quality, and were able to cite substantive differences.

Research question two. *Is the evaluation instrument a useful tool for local educators to assess the quality of their CTE programs?* The evaluation instrument was evaluated through the feedback questionnaire in Phase One and the interviews in Phase Two for its utility, including ease of use, usefulness of results and practicality across programs.

Related to ease of use, participants ranked four criteria on a five-point Likert scale. The criteria included instructions for use, number of criteria, rating scale provided, and time to complete. Overall, each element was rated on average at least "good," or a 4.0 on the five-point scale, with most being between "good" and "very good." Respondents were also asked how easy or difficult it was to find the information to complete the evaluation, and the majority responded positively, with 38.24% ($n=13$) responding "very easy" and 41.18% ($n=14$) responding "easy."

During the Phase Two interviews, it was noted that administrators and educators may have different perspectives on their programs' quality, and different levels of expertise. Educators, generally classroom-based instructors or teachers, at times lack knowledge of national trends and best practices in CTE. In turn, administrators, such as district-level CTE staff or school principals, at times lack knowledge on the details of individual programs. In this regard, it was suggested that additional resources and/or instructions be provided to help these two different populations use the tool effectively.

Related to usefulness, respondents were asked whether they gained new insights into improving their program from completing the instrument, and 82.35% responded "yes" to this question. Comments suggested that participants picked up specific ideas for how to improve

their programs through the evaluation process, and engaged in conversations they might not otherwise have had about program elements. A small group who did not learn much from completing the instrument actually reinforced the usefulness from a different perspective—their programs were already recognized as high quality and thus the results were not surprising. While high-quality programs may not learn as much about their programs from the instrument, it can still serve as reinforcement and provide further validation of their work.

In addition, respondents rated the tool on its usefulness to evaluate and improve program quality. Over half, or 55.88% ($n=19$) said the instrument was “very useful,” and an additional 38.24% ($n=13$) rated it “moderately useful.” When asked how the instrument should be used in the future, teacher professional development was the response selected most often, by 27 (81.82%) respondents. The next most common responses were highlighting best practices and technical assistance to local programs, each with 23 (69.70%) responses.

The usefulness of the tool was another theme that emerged in the interviews in Phase Two, due to a gap in currently available evaluation tools, the depth of information in this evaluation instrument, and its national significance. Across the board, participants agreed the evaluation instrument was an effective program improvement tool. The instrument was described as “a valuable process,” “an enhancement to what we already do,” and useful for “professional growth.” Another identified the tool’s results as “a catalyst” for further discussion and change, and one “appreciated the opportunity to compare” results across programs. Participants cited a key benefit of the tool as drilling down into programs and identifying those key areas that needed additional focus or improvement. It was also noted that results on such a tool, research-based and endorsed nationally, would provide useful credibility and recognition for programs.

Finally, related to practicality across programs, only about one-third, or 32.35%, suggested there were elements that did not apply to their program, and about that same number, 33.33% suggested additional elements that would be helpful to include. A few participants provided comments that suggested improving the online interface, including an opportunity for respondents to provide more detail or evidence, and altering a few criteria for different education levels (middle school, secondary, postsecondary). Despite all offering different programs in different settings, interview participants echoed the idea that the tool was practical across settings, with several mentioning that they felt more programs at their school or in their area could benefit.

Discussion and Conclusions

Prior research has shown that there is a gap in program quality frameworks at the national level that are applicable to individual, local programs (Imperatore & Hyslop, 2015a). Many of the frameworks or evaluations available either do not apply to local programs, or do not provide evidence of a strong research base linked to student outcomes. It was with this backdrop that ACTE set out to develop its High-quality CTE Framework to address concerns that without high-quality program evaluation metrics, it was becoming increasingly difficult for CTE programs to justify their existence (Duncan, 2011).

The conceptual framework of this study relied on the principles of theory driven evaluation and criterion validity. Program-theory driven evaluation is based on underlying concept that a program's elements lead to its outcomes (Coryn, Noakes, Westine, & Schroter, 2010), or that inputs lead to outcomes. The extensive development process leading to the ACTE Framework focused on program inputs—with the aim of identifying the inputs that would have the largest impact on student outcomes.

Results on the instrument. While not a specific research question, the overall results or “scores” on the evaluation instrument provide some interesting information for the program quality conversation, and underlie the conclusions on validity and utility. First, it is important to note that the range of scores (145 to 300) submitted was quite wide, and the mean was well below what would be considered “perfect” on the instrument. This is a possible indication that those submitting program-level data did so in a thoughtful way, and were not merely seeking to make their programs look successful. It was also interesting and enlightening that the “Data and Program Improvement” element was the lowest scoring of all 12 elements. This element includes criteria related to program evaluation and the access to and use of data. Considering the trouble that many participants had accessing their Perkins performance data, and the large gaps identified during the study related to program evaluation, it is no surprise that more attention is needed at the local level on this topic. For program evaluation instruments like this one to have a real impact, there must be a renewed focus on the need for data-driven program improvement processes at the local level.

Validity of evaluation instrument. Within this study, the purpose was to determine if the CTE program inputs identified in the ACTE Framework could be linked to CTE program outcomes as measured by widely accepted federal reporting measures.

Relatedly, criterion-related validity is used to determine the strength of that linkage. Criterion related validity is “the degree of correspondence between a test measure and one or more external referents (criteria), usually measured by their correlation” (Drost, 2011, p. 118). Using principles of criterion-related validity, it was hypothesized that CTE programs that scored higher on the evaluation instrument would have higher levels of student achievement as measured by Perkins performance indicators for technical skill attainment, placement and

credential attainment, and programs that scored lower overall on the evaluation instrument would have lower outcomes on those student achievement measures.

The data collected in the study moderately support this validity-related hypothesis. The sample size was low, and while a positive relationship was found between the overall score on the evaluation instrument and each of the three measures of student performance, it was only statistically significant at the $p < .10$ level in two of the three cases, for technical skill attainment and completion as reported under Perkins performance indicators. Correlation results were not significant for the placement indicator, which could be due to a number of factors outside of the instrument, including both data quality issues and the impact of economic conditions on placement rates. Regarding face validity, additional insight collected from participants suggested that the instrument is indeed designed to gauge key elements associated with program quality.

Based on these results, it can be concluded that there is some evidence of instrument validity, but further refinement and more data and research are needed to ultimately determine how well the instrument reflects the necessary elements of program quality leading to improvement in student outcomes. Since there has been no research published on earlier quality documents and their impacts on student achievement, it is not possible to compare results with other tools. Nevertheless, the level of validity evidence obtained does provide an introductory level of assurance that using the tool is likely to be helpful in identifying areas that will improve program performance. Essentially, there is enough preliminary evidence here that warrants continued use and research as described later in this chapter.

Utility of evaluation instrument. Results related to the utility of the instrument, including for ease of use, practicality across programs, and usefulness of information, were all

extremely positive, which increases the possibility of wide scale future use. Participants reported gaining valuable knowledge about the program merely by using the evaluation instrument—and that the results would be useful for a wide variety of purposes, from shaping professional development to determining where more resources were needed. Results of the study also showed that the instrument could be used in a variety of different settings, including high schools, community colleges and area CTE centers, and that it was flexibility to accommodate many different program models and content areas, allowing the instrument to be scaled up across the country. This is a feature that was not necessarily present in existing program quality frameworks, many of which were focused on specific models or levels of education (Imperatore & Hyslop, 2015a). While the low number of participants could be viewed as a negative reflection on the instrument’s utility, the issues with participation were primarily around the availability of the required Perkins data. Many individuals that expressed interest in participating were not able to due to the data requirements of the study, which were not a part of the underlying evaluation instrument on its own.

In general, this study provided evidence that the ACTE evaluation instrument represents a promising tool to help fill the gap identified in current efforts, although further study may be needed. It also reinforces the call for CTE to do more to evaluate programs in a way that not only identifies positive student outcomes, but also allows the practices that led to these outcomes to be more easily recognized and such information to be more easily shared across the country—opening the “black box” (Suchman 1967) to determine how programs produce results. The element subscores from this instrument have the potential to contribute to this identification by highlighting programs performing extremely well in certain areas. This study also provided an

important contribution to the continued refinement of the instrument to further increase its utility and address remaining concerns.

Overall, this study was a significant step in realizing the potential of this instrument to help determine and improve the quality of CTE programs for millions of students around the country. The CTE community is primed and ready to tackle difficult quality-related issues through the new Perkins Act, but needs tools and resources to support those efforts, and this program evaluation instrument can be one of those resources.

Limitations

There were a number of limitations to the results of this research. Chief overall was the smaller than hoped-for sample size, which limited the statistical power of the correlation results and limited the ability to examine data disaggregated by different program characteristics. For example, there were not enough postsecondary programs to separate out and run valid data on those programs alone. The smaller sample size also limits the ability to generalize results more broadly across the CTE system.

Related to the above limitation, the fact that Perkins program-level data was much harder for programs to obtain than expected also is likely to have changed the nature of participants. Several individuals reported “giving up” on obtaining data and did not submit that program information. Due to the initial length of the instrument, and the challenges that many had to overcome, it is probable that only the most dedicated and interested individuals completed all parts of the study. Positive results, especially related to how participants felt about the utility of the instrument, could have been skewed by this factor. Just by choosing to participate, participants were indicating that they liked the project and wanted it to be successful.

The quality of Perkins data itself was also a limitation of the project's validity results. Different states and even in some cases, different school districts or postsecondary institutions, have different definitions for students, known as "CTE concentrators" included in their Perkins data system, and/or different definitions of CTE programs. This means that in some cases, the data are not truly comparable across states. For the purposes of this study, similar postsecondary and secondary measurement data were also combined, and at times these data points are measured differently as well. For example, technical skill attainment might have been measured by an assessment at the secondary level, but by course grades at the postsecondary level in some programs. While the underlying constructs are the same, having more assurances that the data represented the same students, same "dose" of CTE, and same measurement approaches would have made the study stronger. Some of these changes, such as common CTE concentrator definition, will be made when the new federal CTE legislation goes into effect in 2019 (Advance CTE & ACTE, 2018), presenting opportunities for future, more rigorous research.

Another limitation of the study was related to the inability to control for confounding variables that could have impacted performance on the Perkins indicators. Prior research has established that there are many individual student demographic characteristics that impact educational outcomes (Bowen, Chingos, & McPherson, 2009). These variables were not included in the statistical analysis, and it is unknown the magnitude of the populations served by the programs had on the results.

Since the data in the study were all self-reported, this also introduces limitations. Participants in the interviews specifically raised the idea that educators were likely to score their programs differently than administrators, and the study included a mix of both educators and administrators reporting on their programs. Even within each group, the overall knowledge and

preconceived notions about program quality within each participant could have affected scores on the evaluation instrument.

Finally, the fact that both the researcher and organization, ACTE, were personally known to most participants is a limitation of the study. Participants may have responded to questions more positively than they would have with an external, unknown researcher, particularly during the interviews in the second phase of the study. While efforts were made to choose individuals who would share openly and honestly in this phase of the project, the impact of the personal relationship still could have been a factor in responses.

Next Steps in ACTE's High-quality CTE Initiative

Given the positive results for the study related to the evaluation instrument's utility, and moderately positive results related to its validity, there are a number of next steps within ACTE's high-quality CTE initiative. The first step in the process is revising the underlying Framework based on comments made during the evaluation. Most of the changes suggested were minor, including things like ensuring all the language is appropriate across all levels of education, reviewing individual criteria to ensure that each only evaluates a single construct, reducing any duplication across the Framework, and addressing the specific element of CTSOs. After these adjustments are made, a new version of the underlying Framework will be released.

The next step will be making the evaluation instrument accessible to the broader CTE community. While the Survey Monkey platform was useful for this research, participants suggested a more user-friendly format for the evaluation instrument. Based on feedback obtained during the study, both a new print version and online tool will be created. The online tool will allow participants to receive immediate feedback on their overall score on the instrument and subscores on each element in order to identify the areas most in need of

improvement, which was part of a key theme that emerged during the interviews. It will also allow immediate access to definitions of terms. Results from this version will be able to be saved or printed. The new print version of the form will also walk evaluators through totaling scores and identifying key areas of focus. Later versions will also attempt to address another comment received during the pilot testing, which was to add a place for elaboration or evidence to be attached to results.

Making the evaluation instrument available to the entire CTE community has a number of implications for next steps in professional development and program recognition as well. In fact, shaping professional development activities was the top area suggested by pilot participants for future use of the instrument. By examining areas most often cited as in need of improvement, professional development both nationally and within states or local communities can be more closely tailored to the needs of practitioners. In depth-trainings, conference sessions, articles, and online resources can all be developed and organized to better align with the elements where program need additional attention to improve.

Programs that score well on the instrument as a whole or on one particular element can also be highlighted as “best practices,” with the research backing to validate that claim (Robinson, Shore, & Enersen, 2007). This has implications for program recognition, and helps to address the criticisms of CTE programs’ failure to share and grow evidence-based practices that help all programs produce more desired outcomes (U.S. Department of Education, Office of Vocational and Adult Education, 2012).

While most pilot participants were not interested in using the tool for high stakes decisions around funding, particularly without more external validation of results, it did appear that participants were receptive to it as a part of state or local program evaluation purposes,

particularly if those activities were focused on technical assistance and program improvement and not punitive consequences. Along with this, the new Strengthening Career and Technical Education for the 21st Century Act will require school districts and postsecondary institutions to conduct a local needs assessment, which must include evaluating the quality and implementation of their CTE programs in order to determine where to focus their limited Perkins resources. The evaluation instrument, based on feedback received in the pilot, seems like a natural fit for that activity, and additional tools and resources for how to use it for that purpose will be developed moving forward.

Other companion tools and resources, such as those further elaborating on individual elements, will be developed as well, and should be informed by some of the future research discussed below.

Implications for Policy

Two clear implications for public policy have emerged from this study. The first is that more must be done to make program and institution-level data available to CTE practitioners in order to drive program improvement activities. The difficulty that individual teachers and administrators had accessing data on their own local programs presents challenges far beyond the scope of this study. Without access to this local data, educators have no way of knowing which programs are performing well and which need more attention, or ultimately how to best meet the needs of the students they serve. State and federal policymakers should take a deep look at the functionality of state longitudinal data systems and other data collection and reporting system, and begin to require that local data be available to local educators.

While this outcome data is essential, program evaluation must go further. Participants, especially in the Phase Two interviews, expressed frustration at having to use evaluation tools

that were ill-suited for CTE programs or not comprehensive. Policymakers should require the use of CTE-specific instruments to evaluate CTE programs and educators, and provide the resources necessary for states and local providers to either develop their own unique tools or modify and fully utilize tools available in the field, such as ACTE's evaluation instrument.

Implications for Future Research

Due to the limitations mentioned above, particularly the small sample size and Perkins data issues, further research is needed to more definitely answer questions of validity around the evaluation instrument. Next steps for research in this area might include a larger data collection effort once the online interface is complete, which will ease the user experience. This additional validity research could also be aided by changes in the Perkins data system that will go into effect in 2019, including a provision directing states and local funding recipients to disaggregate their performance data by program of study. This federal directive should begin to move the needle on the availability of program-level data. The new Perkins Act will also require different performance measures that are also more standard across secondary and postsecondary education, leading to higher quality data.

A larger sample and better Perkins data would also allow for other relationships to be examined in more depth, such as individual element subscores or criteria scores and individual performance metrics. For example, some states will collect work-based learning data under the new Act. This could be directly correlated with subscores on the work-based learning element to address validity. They would also allow for more in-depth questions about program performance, such as which elements impact specific student outcomes the most.

It would also be useful to dive more in-depth into the interview questions asked to those who completed the evaluation on two programs—examining the validity of the instrument across

multiple programs at the same location, where student populations, data and outcome measures would be more standard. Evaluating program results across an entire school district or postsecondary institution could shed valuable light not only on the validity and utility of the instrument, but on the programs themselves that could be then benchmarked against the Framework and each other.

In addition, additional research on the specific results of the evaluation instrument would be helpful to those within CTE working to improve practice. Questions like:

- Which elements and criteria are the highest and lowest scoring?
- Do those answers vary based on education level? Program setting? Subject area?
- What resources are needed to help address low-performing areas?
- Are there criteria missing that play an important role in program quality?

In addition, attention on “high-quality CTE” has only increased since ACTE launched the original project and study, and new organizations and experts have weighed in on elements around program evaluation. While some of these new additions were used during the various revisions of the Framework, it would be useful to replicate the first step in ACTE’s original research process. A new content analysis of and crosswalk between quality CTE frameworks including all of the new publications that have been released in the last few years and are likely to be released in the coming months would be a useful addition to the research around quality CTE, and the ACTE Framework could be included.

Finally, a long-term goal of ACTE’s related to this instrument is the development of an awards program or other recognition program related to high-quality CTE programs. Eventually, it would be helpful to establish a “cut score” on the instrument or individual elements that would

help to identify “high-quality” programs. The sample size in this study was not large enough to address this important question, but it should be a key focus of future research.

REFERENCES

- Advance CTE. (2016). *Career clusters*. Retrieved from <https://careertech.org/career-clusters>
- Advance CTE & Association for Career and Technical Education. (2018). *Legislative summary and analysis: Strengthening Career and Technical Education for the 21st Century Act*. Retrieved from https://www.acteonline.org/wp-content/uploads/2018/08/AdvanceCTE_ACTE_P.L.115-224Summary_Updated080618.pdf
- Aliaga, O. (2012, February). *Vanguard practices in CTE: A podcast with Oscar Aliaga* [Audio Podcast]. Retrieved from <http://www.nrccte.org/resources/podcasts/vanguard-practices-cte-podcast-oscar-aliaga>
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education, & Joint Committee on Standards for Educational and Psychological Testing (U.S.). (2014). *Standards for educational and psychological testing*.
- Association for Career and Technical Education. (n.d.). *CTE today!* Retrieved from <https://www.acteonline.org/WorkArea/DownloadAsset.aspx?id=1909>
- Association for Career and Technical Education. (2006). *Perkins Act of 2006: The official guide*. Alexandria, VA: Author.
- Association for Career and Technical Education & Advance CTE. (2016). *State policies impacting CTE: 2015 year in review*. Retrieved from [http://www.acteonline.org/uploadedFiles/Who_We_Are/Press/2015_State-Policy-Review_FINAL%20\(1\).pdf](http://www.acteonline.org/uploadedFiles/Who_We_Are/Press/2015_State-Policy-Review_FINAL%20(1).pdf)
- Barlow, M. L. (1976, May). The vocational age emerges, 1876-1926. *American Vocational Journal*, 46-58.
- Bowen, W. G., Chingos, M. M., & McPherson, M. S. (2009). *Crossing the finish line: Completing college at America's public universities*. Princeton, NJ: Princeton University Press.
- Braun, V., & Clarke, V. (2006) Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101.
- Carl D. Perkins Career and Technical Education Act of 2006, 20 U.S.C. § 2301 *et seq.* (2012)

- Coryn, C. L. S., Noakes, L. A., Westine, C. D., & Schoter, D. C. (2011). A systematic review of theory-driven evaluation practice from 1990 to 2009. *American Journal of Evaluation*, 32(2), 199-226.
- Davidson, E. J. (2000). Ascertaining causality in theory-based evaluation. In P. J. Rogers, T. A. Hacsí, A. Petrosino, & T. A. Huebner (Eds.), *Program theory in evaluation: Challenges and opportunities. New Directions for Evaluation* (No. 87). San Francisco, CA: Jossey-Bass.
- Donaldson, S. I. (2007). *Program theory-driven evaluation science: Strategies and applications*. New York, NY: Routledge.
- Drost, E. A. (2011). Validity and reliability in social science research. *Education Research and Perspectives*, 38(1), 105-123.
- Duncan, A. (2011, April). *Rigor, relevance, and the future of career and technical education*. Speech presented at the meeting of the National Association of State Directors of Career Technical Education Consortium, Washington, DC. Retrieved from U.S. Department of Education website: <http://www.ed.gov/news/speeches/rigor-relevance-and-future-career-and-technical-education>
- Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2011). *Program evaluation: Alternative approaches and practical guidelines*. Upper Saddle River, NJ: Pearson Education, Inc.
- Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational research: An introduction* (8th ed.). Boston, MA: Pearson Education, Inc.
- Gordon, H. R. D. (2014). *The history and growth of career and technical education in America* (4th ed.). Long Grove, IL: Waveland Press, Inc.
- Grimm, L. G. (1993). *Statistical applications for the behavioral sciences*. New York, NY: John Wiley & Sons, Inc.
- Hayward, G. C., & Benson, C. S. (1993). *Vocational- technical education: Major reforms and debates, 1917- present*. Washington, DC: U.S. Department of Education, Office of Vocational and Adult Education.
- Hess, F. M., & Little, B. (2015). *“Moneyball” for education”: Using data, evidence, and evaluation to improve federal education policy*. Washington, DC: American Enterprise Institute and Results for America.
- Hollenbeck, K. M. (2011). *Conducting return on investment analyses for secondary and postsecondary CTE: A framework*. Louisville, KY: University of Louisville, National Research Center for Career and Technical Education.

- Hyslop, A. (2015, October 26). *Senate kicks off its Perkins reauthorization efforts* [Web log post]. Retrieved from <http://ctepolicywatch.acteonline.org/2015/10/senate-kicks-off-its-perkins-reauthorization-efforts.html>
- Imperatore, C., & Hyslop, A. (2015a). *Defining high-quality CTE: Contemporary perspectives on CTE quality*. Alexandria, VA: Association for Career and Technical Education.
- Imperatore, C., & Hyslop, A. (2015b). *Defining high-quality CTE: Quality CTE program of study framework, Draft version 1.0*. Alexandria, VA: Association for Career and Technical Education.
- Imperatore, C., & Hyslop, A. (2016). *Defining high-quality CTE: Quality CTE program of study framework, Draft version 3.0*. Alexandria, VA: Association for Career and Technical Education.
- Imperatore, C., & Hyslop, A. (2017). *Defining high-quality CTE: Quality CTE program of study framework, Beta version 4.0*. Alexandria, VA: Association for Career and Technical Education.
- Kirkpatrick, D. L., & Kirkpatrick, J. D. (2006). *Evaluating training programs*. San Francisco, CA: Berrett-Koehler Publishers, Inc.
- Kirkpatrick, J. D., & Kirkpatrick, W. K. (2009). *Kirkpatrick then and now: A strong foundation for the future*. St. Louis, MO: Kirkpatrick Publishing.
- Klein, S., Sheil, A. R., White, R., Staklis, S., Alfeld, C. Dailey, C. R., ...Poliakoff, A. (2014). *Evaluation of the implementation of the Carl D. Perkins Career and Technical Education Act of 2006*. Retrieved from RTI International website <https://www.rti.org/sites/default/files/resources/evaluationofimplementationofperkinsactof2006.pdf>
- LaFollette, A. M. (2011). *An historical policy analysis of the Carl D. Perkins legislation: Examining the history, creation, implementation, and reauthorization of the law* (Doctoral dissertation). Retrieved from https://www.ideals.illinois.edu/bitstream/handle/2142/26060/LaFollette_Aimee.pdf?sequence=1&isAllowed=y
- Linked Learning. (2012). *Criteria for Linked Learning pathway quality review and continuous improvement*. Retrieved from http://www.connectedcalifornia.org/direct/files/certification/Certification_Rubric_Booklet_121112_secure.pdf
- Lintner, L., & Haigh, J. (2013, June). *Perkins and evaluation of CTE programs: Sources and content*. Session presented at the Virtual Data Quality Institute of the U.S. Department of Education, Office of Vocational and Adult Education, Division of Academic and Technical Education. Retrieved from

http://cte.ed.gov/dqi/app/webroot/files/S3_2013_DQI_CTE_Program_Evaluation_Lintner_Haigh.pdf

- Locke, L., Spirdusa, W., & Silverman, S. (2007). *Proposals that work: A guide for planning dissertations and grant proposals*. Thousand Oaks, CA: Sage Publications, Inc.
- Lynch, R. L. (2000). *New directions for high school Career and Technical Education in the 21st century*. Columbus, OH: ERIC Clearinghouse on Adult, Career, and Vocational Education, Center on Education and Training for Employment, College of Education, The Ohio State University.
- Missouri Department of Elementary and Secondary Education, Office of College and Career Readiness. (2015). *Common criteria and quality indicators for career education programs*. Retrieved from https://dese.mo.gov/sites/default/files/cte_common_criteria_and_quality_indicators_march2015.pdf
- MPR Associates, Inc. (2010). *Programs of study: Local implementation readiness and capacity self-assessment: A tool for local college and career readiness*. Berkeley, CA: MPR Associates, Inc., under contract from the U.S. Department of Education.
- National Career Academy Coalition. (2013). *National standards of practice*. Retrieved from http://www.ncacinc.com/index.php?option=com_content&view=article&id=15&Itemid=10
- National Dissemination Center for Career and Technical Education. (2002). *Exemplary career and technical education programs: Secondary-level user's guide to program improvement*. Retrieved from <http://www.azed.gov/career-technical-education/files/2011/06/exemplaryprogramrubric.pdf>
- National Research Center for Career and Technical Education. (2012). *Vanguard practices in CTE*. Retrieved from <http://www.nrccte.org/resources/studies/vanguard-practices-cte>
- Office of Vocational and Adult Education. (2010). *Career and technical programs of study: A design framework*. Retrieved from http://cte.ed.gov/nationalinitiatives/POS_Framework_Components_and_Subcomponents_2-24-10.doc
- Robinson, A., Shore, B., & Enerson, D. (2007). *Best practices in gifted education: An evidenced-based guide*. Waco, TX: Prufrock Press Inc.
- Rogers, P. J. (2000). Program theory evaluation: Not whether programs work but how they work. In D. L. Stufflebeam, G. F. Madaus, & T. Kellaghan (Eds.), *Evaluation models: Viewpoints on educational and human services evaluation* (pp. 209-232). Boston, MA: Kluwer.

- Ruffing, K. (2006). *A brief history of Career Clusters: Supporting a new vision for CTE*. Retrieved from <http://ocrl.illinois.edu/docs/librariesprovider4/pos/careerclusterhistory.doc?sfvrsn=4>
- Scriven, M. (1991). *Evaluation thesaurus* (4th ed.). Newbury Park, CA: Sage.
- Silverberg, M., Warner, E., Fong, M., & Goodwin, D. (2004). *National assessment of vocational education: Final report to Congress*. Washington, DC: U.S. Department of Education, Office of the Under Secretary, Policy and Program Studies Service.
- Stone, J. R. III, & Struebing, J. (2008, November). *Tapping the wisdom of practitioners*. Presentation made at the annual meeting of the Association for Career and Technical Education, Charlotte, NC.
- Strengthening Career and Technical Education for the 21st Century Act, H.R. 5587, 114th Cong. (2016)
- Suchman, E. (1967). *Evaluative research: Principles and practice in public service and social programs*. New York, NY: Russell Sage Foundation.
- Tashakkori, A. & Teddlie, C. (Eds.). (2003). *Handbook of mixed methods in social and behavioral research*. Thousand Oaks, CA: Sage Publications, Inc.
- Terrell, S. (2011). Mixed-methods research methodologies. *The Qualitative Report*, 17(1), 254-280. Retrieved from <http://www.nova.edu/ssss/QR/QR17-1/terrell.pdf>
- Trochim, W. M. (2006). *The research methods knowledge base, 2nd edition*. Retrieved from <http://www.socialresearchmethods.net/kb/>
- U.S. Department of Education. (1983). *A nation at risk*. Retrieved from <https://www2.ed.gov/pubs/NatAtRisk/risk.html>
- U.S. Department of Education (2013). "Program description." *Investing in Innovation Fund (i3)*. Retrieved from <http://www2.ed.gov/programs/innovation/index.html>
- U.S. Department of Education, Office of Career, Technical, and Adult Education. (2015, April). *The evolution and potential of career pathways*. Washington, DC: Author.
- U.S. Department of Education, Office of Career, Technical, and Adult Education, Division of Academic and Technical Education (n.d.a) Accountability. *Perkins Collaborative Resource Network*. Retrieved from <http://cte.ed.gov/accountability/accountability-for-results>

- U.S. Department of Education, Office of Career, Technical, and Adult Education, Division of Academic and Technical Education (n.d.b) Programs of study: Promoting rigorous programs of study (RPOS) project. *Perkins Collaborative Resource Network*. Retrieved from <http://cte.ed.gov/initiatives/promoting-rigorous-programs-of-study-rpos-project>
- U.S. Department of Education, Office of Planning, Evaluation and Policy Development, Policy and Program Studies Service (2014), National Assessment of Career and Technical Education: Final Report to Congress. Washington, DC.
- U.S. Department of Education, National Center for Education Statistics. (2009). *High school transcript study (Tables H122-124)*. Retrieved from <http://nces.ed.gov/surveys/ctes/tables/index.asp?LEVEL=SECONDARY>
- U.S. Department of Education, Office of Vocational and Adult Education. (2012, April). *Investing in America's future: A blueprint for transforming career and technical education*. Washington, DC: Author.
- U.S. Departments of Education, Labor, and Health and Human Services. (2012, April 4). Joint letter. Retrieved from <https://www2.ed.gov/about/offices/list/ovae/ten-attachment.pdf>
- Valentine, J. C., Hirschy, A. S., Bremer, C., Novillo, W., Castellano, M., & Banister, A. (2009). *Systematic reviews of research: Postsecondary transitions - Identifying effective models and practices*. Louisville, KY: National Research Center for Career and Technical Education.

APPENDICES

Appendix A: Institutional Review Board (IRB) Letter



RESEARCH INTEGRITY AND COMPLIANCE
Institutional Review Boards, FWA No. 00001669
12901 Bruce B. Downs Blvd., MDC035 • Tampa, FL 33612-4799
(813) 974-5638 • FAX(813)974-7091

November 22, 2017

Alisha Hyslop
L-CACHE - Leadership, Counseling, Adult, Career & Higher Education
Tampa, FL 33626

RE: **Not Human Subjects Research Determination**

IRB#: Pro00033073

Title: Testing the Validity and Utility of a Career and Technical Education Programs of Study
Evaluation Instrument

Dear Ms. Hyslop:

The Institutional Review Board (IRB) has reviewed your application. The activities presented in the application involve methods of program evaluation, quality improvement, and/or needs analysis. While potentially informative to others outside of the university community, study results would not appear to contribute to generalizable knowledge. As such, the activities do not meet the definition of human subject research under USF IRB policy, and USF IRB approval and oversight are therefore not required.

While not requiring USF IRB approval and oversight, your study activities should be conducted in a manner that is consistent with the ethical principles of your profession. If the scope of your project changes in the future, please contact the IRB for further guidance.

If you will be obtaining consent to conduct your study activities, please remove any references to "research" and do not include the assigned Protocol Number or USF IRB contact information.

If your study activities involve collection or use of health information, please note that there may be requirements under the HIPAA Privacy Rule that apply. For further information, please contact a HIPAA Program administrator at (813) 974-5638.

Sincerely,

A handwritten signature in black ink, appearing to read "Kristen Salomon", with a horizontal line extending to the right.

Kristen Salomon, Ph.D., Vice Chairperson
USF Institutional Review Board

Appendix B: Conflict of Interest Disclosure Approval and Management Plan



MEMORANDUM

TO: Alisha Hyslop
FROM: Amber Moore, COI Manager
RE: Approval/Return of Financial Relationships Disclosure for COI Review of Disclosure for Testing the Validity and Utility of a CTE Program of Study Evaluation Instrument
DATE: 11/30/2017

On 11/28/2017 the Conflict of Interest Committee reviewed and approved your eCOI Disclosure in the above referenced study.

The Committee's approval is based on the amount and nature of the interest you have disclosed. Should your financial interests change during the course of research, please advise the Committee by updating your eCOI Disclosure at <https://arc.research.usf.edu/prod>.

The Committee's approval is further contingent on the continued veracity of the facts that provide the basis for the approved management plan. Alteration of the study design on which the management plan is based or failure to adhere to the terms of the management plan (if any additional requirements have been imposed) may result in withdrawal of the Committee's approval and possible referral to other committees and/or offices (e.g., Faculty Affairs, Dean's Office, Department Chair, IRB, etc.) for appropriate remedial action.

If you have any questions, please contact Amber Moore, COI Manager, at 974-9345 or amber1@usf.edu.

Final Conflict of Interest Management Plan

Alisha Hyslop

CTE Program of Study Evaluation Instrument Pilot

Status:

- ✓ **Approved** by COI Committee on November 28, 2017.

Involvement of Human Subjects:

The study does not involve human subjects.

IRB: USF IRB Pro 33073. The USF IRB determined that the study qualifies as Not Human Subjects Research (NHSR) as per USF HRPP Policy because it is program evaluation.

Purpose of the Study:

The growing proliferation of career and technical education (CTE)-related standards and frameworks has left the field without a clear definition of program quality, and opens the door for criticisms from federal policymakers related to a lack of consistent quality. The purpose of this study will be to build on the research already completed by Association for Career and Technical Education (ACTE) in the development of the ACTE High-quality CTE Program of Study Framework and accompanying evaluation instrument to examine the instrument's validity and utility when used independently by local educators in a pilot test.

Primary Objectives:

The purpose of this study is to conduct a pilot test to examine the ACTE High-quality CTE Program of Study Evaluation Instrument's validity and utility when used independently by local educators. The instrument will be tested both for its criterion validity—whether the included elements are appropriate measures of program quality, as related to already established program outcome measures—as well as its utility, including ease of use, practicality, and usefulness of results to local practitioners. In addition, the resources, barriers and possibilities for scaling up the tool to widespread use will be examined. The specific research questions that will be examined to achieve the purposes of the study are as follows:

1. Does the evaluation instrument provide valid evidence of program quality?
2. Is the evaluation instrument a useful tool for local educators to assess the quality of their CTE programs?

Study Procedures:

This research study will be conducted as a two-phase, sequential mixed-method study. During Phase One, participants (either teachers or administrators volunteers) will be asked to complete the Evaluation Instrument on one of the career and technical education programs they work with. They will then be asked to complete the Evaluation Instrument Feedback Questionnaire to submit thoughts about the experience of using the questionnaire and its value, as well as aggregate program performance data that can be compared to evaluation results. In Phase Two, five participants will be selected for more in-depth interviews to gain deeper insights into the utility of the instrument.

The sample for the first phase of the study will consist of a minimum of 30 individual CTE programs of study. Each program will be represented in the study by a key contact, either a CTE teacher or administrator, who agrees to participate in the research and complete required data collection instruments. Participants will be selected primarily through a non-probability convenience sample, with volunteers considered first. Throughout the foundational research that has led to this study, a list of potential volunteers has been collected. These individuals will be contacted and given the first opportunity to participate in the pilot of the instrument. To aid with the issue of generalizability, the sample will then be filled out using purposive sampling techniques to achieve geographic diversity and diversity of CTE program areas, using the membership of ACTE as the sampling frame. The sample for the second phase of the study will consist of five individuals sampled purposefully from the Phase One participants for one-on-one interviews with the researcher. Criteria for selection will include diversity of scores on evaluation instrument, geographic diversity, CTE program diversity, and educational setting diversity, with an attempt made to include programs in high schools, area CTE centers and postsecondary institutions. No personal information will be collected on the individuals completing the evaluation and feedback questionnaire—only CTE program information, and data will not be released in any way that is identifiable.

Role of Investigator in the Study:

Alisha Hyslop is the PI of the study and therefore oversees all aspects of the study. She participated in the study design and will conduct study initiation activities, perform experimental portions of the study, record raw data, analyze or interpret data, record/evaluate/present study results and write the final report.

Outside Entities Involved in the Research:

The study is non-sponsored.

Amount and Nature of Interest:

Alisha Hyslop is employed full time as the director of public policy at the Association for Career and Technical Education (ACTE) for the last 15+ years. This study will pilot test the ACTE High-quality CTE Program of Study Framework and accompanying evaluation instrument. Alisha co-authored the framework and co-developed the evaluation instrument.

Management Plan:

The following inherent qualities of the study protect against or neutralize bias on the part of an investigator:

1. The study will collect participant-reported data and all interviews will be transcribed.
2. The study involves four dissertation committee members (Dr. Victor Hernandez, Dr. Liliana Rodriguez, Dr. Edward Fletcher and Dr. Robert Dedrick, College of Education) who do not have a related interest.

The following management controls may also protect against bias on the part of the investigator:

3. Alisha Hyslop will disclose the interest to research personnel and in resulting publications and presentations.

Referred by:

Investigator

USF Conflict of Interest Committee Action:

Management Plan approved (with or without modifications)?

✓Yes.

Summary of Discussion: The COI Committee considered the nature of the project and the inherent controls including participant-reported data, transcribed interviews and four non-interested dissertation committee members. Therefore, the Committee determined the study was well-controlled with the additional control that Alisha Hyslop will disclose the interest to research personnel and in resulting publications and presentations.

Modification requested?

✓No.

Monitoring required?

✓No.

Appendix C: Evaluation Instrument Rating Scale

0 = Not at all achieved

No evidence of the criterion in the program of study.

1 = Minimally achieved

The criterion is minimally implemented in the program of study. For example:

- implementation is just beginning
- implementation is evident infrequently
- implementation is evident in a small portion of the program of study
- access is limited to a small segment of students

2 = Moderately achieved

The criterion is evident in the program of study, but implementation is uneven or incomplete. For example:

- only part of the criterion is evident
- implementation is evident part of the time, but not on a sustained and regular basis
- implementation is evident in portions of the program of study
- access is available to most students

3 = Substantially achieved

The criterion has been fully implemented throughout the entire program of study. For example:

- all parts of the criterion are evident
- implementation is evident on a regular and sustained basis
- implementation is evident across all portions of the program of study
- access is available to all students

Appendix D: Evaluation Instrument

Note: This instrument was presented in Survey Monkey, and screenshots are included below. Spacing and other visuals may not be an accurate representation of the online experience.

Part 1: Program of Study Information

This information will be used to match your Evaluation Instrument results with the information submitted on the Feedback Questionnaire, but will not be used publicly in any way.

Name of Program of Study:

School/Institution(s) involved in Program of Study:

Name of individual submitting Evaluation Instrument:

E-mail:

Part 2: Program of Study Evaluation

Please use the categories provided to rate your Program of Study on each of the criteria under the 12 elements of ACTE's High Quality CTE Framework. Below are descriptions of each of the categories to help you choose which best describes your program:

0 = Not at all achieved

No evidence of the criterion in the program of study.

1 = Minimally achieved

The criterion is minimally implemented in the program of study. For example:

- implementation is just beginning
- implementation is evident infrequently
- implementation is evident in a small portion of the program of study
- access is limited to a small segment of students

2 = Moderately achieved

The criterion is evident in the program of study, but implementation is uneven or incomplete. For example:

- only part of the criterion is evident
- implementation is evident part of the time, but not on a sustained and regular basis
- implementation is evident in portions of the program of study
- access is available to most students

3 = Substantially achieved

The criterion has been fully implemented throughout the entire program of study. For example:

- all parts of the criterion are evident
- implementation is evident on a regular and sustained basis
- implementation is evident across all portions of the program of study
- access is available to all students

1. Standards-aligned and Integrated Curriculum

	0 = Not at all achieved	1 = Minimally achieved	2 = Moderately achieved	3 = Substantially achieved
a. The program of study curriculum is developed with employer input to prepare students for both further education and in-demand and emerging careers.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b. The curriculum is based on industry-validated technical standards and competencies.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c. The curriculum is aligned with relevant content and standards for core subjects, such as reading, math and science, including federal, state and/or local standards, as appropriate.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d. The curriculum incorporates employability skill standards that help students succeed in the workplace, such as problem solving, critical thinking, teamwork, communications and workplace etiquette.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e. The curriculum allows for student application of integrated knowledge and skills in authentic scenarios.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
f. Program of study standards are publically available and accessible, as appropriate, to students, parents/guardians, partners and the community.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
g. The curriculum is reviewed regularly by all relevant stakeholders and revised as necessary to reflect the latest advances in the field, evidence-based program models and evaluations of student performance.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

2. Sequencing and Articulation

	0 = Not at all achieved	1 = Minimally achieved	2 = Moderately achieved	3 = Substantially achieved
a. The CTE program of study includes a sequence of courses and/or competencies across secondary and postsecondary education that incorporates technical, academic and employability knowledge and skills.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	0 = Not at all achieved	1 = Minimally achieved	2 = Moderately achieved	3 = Substantially achieved
b. The program of study starts with broad foundational knowledge and skills and progresses in specificity to build students' depth of knowledge and skills.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c. Content and standards within the CTE program of study are non-duplicative and vertically aligned to prepare students to transition seamlessly to the next level of education.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d. The CTE program of study sequence leads to one or more recognized postsecondary credentials, including industry certifications, licenses, apprenticeship certificates, postsecondary certificates and degrees.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e. Students in the CTE program of study have opportunities to earn postsecondary credit while in high school.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
f. Secondary and postsecondary CTE educators, along with representatives of the employer community, collaborate regularly on course sequencing, vertical alignment and opportunities for postsecondary credit.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	0 = Not at all achieved	1 = Minimally achieved	2 = Moderately achieved	3 = Substantially achieved
g. The program of study is coordinated with broader career pathways systems, as defined in the Workforce Innovation and Opportunity Act, such as through the development of stackable credentials, as appropriate and available.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. Student Assessment				
	0 = Not at all achieved	1 = Minimally achieved	2 = Moderately achieved	3 = Substantially achieved
a. Formative and summative assessments are integrated throughout the program of study to validate student learning gains, including both classroom/school-based and standardized, third-party assessments, as appropriate.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b. Assessments are aligned to program standards and curriculum and appropriate to students' current level of knowledge and skill attainment.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c. Assessments are valid, reliable and developed or chosen in accordance with relevant quality standards.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	0 = Not at all achieved	1 = Minimally achieved	2 = Moderately achieved	3 = Substantially achieved
d. The CTE program of study incorporates multiple forms of assessment, including performance-based assessment where students must demonstrate the application of their knowledge and skills.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e. Assessments within the CTE program of study provide objective information on student attainment of industry-validated technical knowledge and skills provided by program curriculum.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
f. Assessments within the CTE program of study provide objective information on student attainment of academic knowledge and skills provided by program curriculum.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
g. Assessments within the CTE program of study provide objective information on student attainment of employability knowledge and skills provided by program curriculum.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
h. The CTE program of study prepares students for assessments that lead to recognized postsecondary credentials, as available and appropriate.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. Prepared and Effective Program Staff				

	0 = Not at all Achieved	1 = Minimally Achieved	2 = Moderately Achieved	3 = Substantially Achieved
a. CTE educators in the program of study meet minimum state, district and/or institution certification and licensing requirements.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b. CTE educators maintain up-to-date knowledge and skills across all aspects of an industry and have appropriate industry-relevant credentials.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c. CTE educators maintain relevant evidence-based pedagogical knowledge and skills.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d. CTE educators engage in ongoing, rigorous professional development on a wide range of topics covering all elements of a high-quality CTE program of study, as described in ACTE's Defining High-quality CTE: Quality CTE Program of Study Framework, which might include pursuit of advanced educator certification.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e. CTE educators demonstrate leadership and commitment to the profession.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
f. CTE educators have the time, resources and supports to implement all elements of a high-quality CTE program of study, as described in ACTE's Defining High-quality CTE: Quality CTE Program of Study Framework.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

0 = Not at all Achieved 1 = Minimally Achieved 2 = Moderately Achieved 3 = Substantially Achieved

g. CTE educators, academic educators, counselors, administrators and other relevant staff collaborate regularly and frequently to coordinate curriculum, instruction, assessment and extended learning activities and to analyze data for program improvement.



5. Engaging Instruction

0 = Not at all Achieved 1 = Minimally Achieved 2 = Moderately Achieved 3 = Substantially Achieved

a. Program of study instruction is driven by relevant content area standards and learning objectives.

b. Project-based learning and related instructional approaches, such as problem-based, inquiry-based and challenge-based learning, are fundamental to the CTE program of study.

c. Contextualized instruction results in students applying technical, academic and employability knowledge and skills within authentic scenarios.

d. Instruction emphasizes the connection between academic and technical knowledge and skills, including through cross-disciplinary collaboration.

e. Instruction incorporates relevant equipment, technology and materials to support learning.

f. Instruction is flexible, differentiated and personalized to meet the needs of a diverse student population.

g. Management of the educational environment builds a culture of learning and respect.

6. Access and Equity

	0 = Not at all Achieved	1 = Minimally Achieved	2 = Moderately Achieved	3 = Substantially Achieved
a. The CTE program of study is promoted to all potential participants and, as appropriate, their parents/guardians, in a manner that is free from bias, inclusive and non-discriminatory.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b. Students who have been traditionally underrepresented, including by gender, race and ethnicity, and special population status, are actively recruited.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c. Career guidance is offered to all potential and current program of study participants in a manner that is free from bias, inclusive and non-discriminatory and that takes into account student interests, preferences and abilities.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d. As appropriate, accommodations to facilities and equipment as well as curriculum, instruction, materials and assessments are provided to ensure all students have the opportunity to achieve success in the program of study.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e. As appropriate, support services, such as tutoring and transportation assistance, are provided to ensure all students have the opportunity to achieve success in the program of study.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	0 = Not at all Achieved	1 = Minimally Achieved	2 = Moderately Achieved	3 = Substantially Achieved
f. Curriculum, instruction, materials and assessments are free from bias, inclusive and non-discriminatory.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
g. Appropriate actions are taken to eliminate barriers to extended learning experiences, such as work-based learning, CTSO participation and postsecondary credit attainment, for all students, including special populations.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7. Facilities and Equipment				
	0 = Not at all Achieved	1 = Minimally Achieved	2 = Moderately Achieved	3 = Substantially Achieved
a. Facilities and equipment used in the program of study reflect current workplace, industry and/or occupational practices and requirements.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b. Facilities and equipment support and align to curriculum standards and program objectives.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c. Facilities and equipment meet appropriate federal, state and local standards for occupational safety and health, as applied in the related industry.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d. Students demonstrate safe and appropriate use and maintenance of facilities and equipment within the CTE program of study.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	0 = Not at all Achieved	1 = Minimally Achieved	2 = Moderately Achieved	3 = Substantially Achieved
e. Processes are defined and resources provided to regularly inspect, update and replace facilities and equipment, as necessary.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
f. Program of study partners collaborate to maximize student access to relevant facilities and equipment.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
g. Relevant materials, tools, supplies and personal protective equipment are available and used appropriately.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
h. Facilities are free from bias, inclusive and non-discriminatory, and they meet all Title IX, Americans with Disabilities Act and other accessibility requirements.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8. Business and Community Partnerships				
	0 = Not at all Achieved	1 = Minimally Achieved	2 = Moderately Achieved	3 = Substantially Achieved
a. Representatives of the program of study actively conduct outreach activities to develop partnerships to ensure the program of study is informed by employer and community needs.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b. The program of study has a formalized, structured approach to coordinating partnerships, such as an advisory board or sector partnership.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	0 = Not at all Achieved	1 = Minimally Achieved	2 = Moderately Achieved	3 = Substantially Achieved
c. Partners include a diverse range of employers; industry representatives; community, workforce and economic development agencies; and other education stakeholders.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d. The program of study has partners who: • identify, validate and review curriculum for technical, academic and employability knowledge and skills that meet the needs of students and industry	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
• identify appropriate assessments and recognized postsecondary credentials	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
• evaluate equipment, facilities and materials to ensure they are consistent with industry standards	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
• provide input on current and future workforce demand and skill needs to inform updates to the program of study	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
• provide input on the further education and training necessary for career pathways	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
• identify, provide and evaluate work-based learning experiences for students	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
• participate in CTSO activities; for example, by serving as mentors and judges	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	0 = Not at all Achieved	1 = Minimally Achieved	2 = Moderately Achieved	3 = Substantially Achieved
<ul style="list-style-type: none"> offer opportunities, such as externships, for educators to stay current with industry-relevant knowledge and skills 	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<ul style="list-style-type: none"> provide support in tangible ways, such as by investing funds, providing in-kind support and/or helping raise external funds to meet program of study goals 	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<ul style="list-style-type: none"> advocate for and promote the program of study 	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<ul style="list-style-type: none"> help to evaluate the effectiveness of the program of study in preparing students for further education and careers 	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9. Career Development				
	0 = Not at all Achieved	1 = Minimally Achieved	2 = Moderately Achieved	3 = Substantially Achieved
<p>a. Comprehensive career development is coordinated and sequenced to promote and support the career decision-making and planning of all students, both prior to entering and during the program of study.</p>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<p>b. Each CTE student in the program of study has a personalized, multi-year education and career plan that reflects exploration of the student's interests, preferences and abilities; and informs course selection, planning for further education and a career, and involvement in extended learning.</p>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	0 = Not at all Achieved	1 = Minimally Achieved	2 = Moderately Achieved	3 = Substantially Achieved
c. Career development activities are aligned with relevant national, state and/or local standards.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d1. Students in the program of study and their parents/guardians, as appropriate, are provided accurate and timely information on extended learning experiences available through the program of study, such as work-based learning, CTSO participation and postsecondary credit attainment.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d2. Students in the program of study and their parents/guardians, as appropriate, are provided accurate and timely information on further education and training options, including application procedures, enrollment, financing, and their projected educational, employment and earnings outcomes.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d3. Students in the program of study and their parents/guardians, as appropriate, are provided accurate and timely information on regional occupational trends and outlook, high-demand and high-wage career opportunities, and the educational pathways that lead to current and projected career opportunities.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	0 = Not at all Achieved	1 = Minimally Achieved	2 = Moderately Achieved	3 = Substantially Achieved
e. Students in the program of study have access to job search information and placement services as they near completion of the program of study.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
f. Guidance, counseling and advisement professionals have access to up-to-date information and training about extended learning experiences, education and training options, and regional occupational trends, in order to aid students in education and career planning and decision making.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

10. Career and Technical Student Organizations (CTSOs)

0 = Not at all Achieved 1 = Minimally Achieved 2 = Moderately Achieved 3 = Substantially Achieved

a. A CTSO is an integral, intra-curricular part of the CTE program of study, available to every student at some point during the program of study.

b. The CTSO is aligned with relevant national, state and/or local standards.

c. CTSO activities develop and reinforce relevant technical, academic and employability knowledge and skills.

d. The CTSO provides opportunities for students to interact with business professionals.

e. The CTSO provides opportunities for students to participate in relevant competitive events.

f. The CTSO provides opportunities for students to participate in relevant community and school service activities.

g. The CTSO provides opportunities for students to participate in leadership development activities.

h. The CTSO is closely supervised by an educator and/or other appropriate staff with clearly defined roles.

11. Work-based Learning

	0 = Not at all Achieved	1 = Minimally Achieved	2 = Moderately Achieved	3 = Substantially Achieved
a. Work-based learning is organized in a sequenced continuum that progresses in intensity as a student moves through the program of study, including a range of activities such as workplace tours, job shadowing, school-based enterprises, internships and apprenticeships.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b. Work-based learning experiences are an integral part of the program of study curriculum and the full continuum is accessible to every student at some point during the program of study.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c. The work-based learning continuum is aligned with relevant national, state and/or local curriculum standards.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d. Work-based learning experiences develop and reinforce relevant technical, academic and employability knowledge and skills.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e. Work-based learning experiences are aligned with each student's multi-year education and career plan.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
f. Work-based learning experiences provide students with meaningful interactions with business professionals at school, in workplaces, in the community and/or virtually.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	0 = Not at all Achieved	1 = Minimally Achieved	2 = Moderately Achieved	3 = Substantially Achieved
g. Requirements and procedures for work-based learning experiences that address access, selection, liability, supervision, safety, transportation, learning objectives and evaluations are formalized with employers, students and, as appropriate, parents/guardians.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
h. Training should be provided to students and employers in advance of work-based learning experiences, addressing each stakeholder's rights and responsibilities as well as safety and appropriate behavior.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
i. Work-based learning experiences comply with relevant federal, state and local laws and regulations.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
j. Work-based learning experiences are closely supervised by an educator and/or other appropriate staff with clearly defined roles.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
k. Employers play a leadership role in developing, facilitating and evaluating work-based learning experiences.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
l. Students engage in reflection and document learning resulting from work-based learning experiences, such as through a portfolio or presentation.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
12. Data and Program Improvement				

	0 = Not at all Achieved	1 = Minimally Achieved	2 = Moderately Achieved	3 = Substantially Achieved
a. Continuous evaluation of each element of ACTE's Defining High-quality CTE: Quality CTE Program of Study Framework, including all stakeholders, supports program improvement.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b. All stakeholders understand why data is collected, how it will be used, and its value in supporting student success.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c. There is a formal process in place for the systematic and continuous use of student performance data for program improvement, including identifying and addressing equity gaps.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d. Collaborative processes and supports are in place in the program of study to ensure the timely and accurate collection and submission of data for required reporting.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e. Federal and state performance indicators form the foundation of data collection, analysis and reporting; however, additional data on student access and performance are included, as appropriate, to aid in program improvement.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
f. Privacy and security protections are in place for data collection, storage, analysis and reporting. Protections should adhere to all federal, state and local privacy laws.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	0 = Not at all Achieved	1 = Minimally Achieved	2 = Moderately Achieved	3 = Substantially Achieved
g. Educators have access to relevant valid and reliable aggregate data on all students participating in the program of study.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
h. Educators have access to relevant valid and reliable data disaggregated by gender, race and ethnicity, and special population status, thus facilitating comparisons of access and performance among subpopulations and with the general student population.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
i. Data is shared, as appropriate, in an easy-to-understand format with students, parents/guardians, partners and the community.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Appendix E: Feedback Questionnaire

Note: This instrument will be presented in Survey Monkey, and screenshots are included below. Spacing and other visuals may not be an accurate representation of the online experience

Copy of HQ CTE Evaluation Feedback Questionnaire

Evaluation Instrument Feedback

1. How well did your responses and scores on the Evaluation Instrument reflect the overall quality of your program?

Very Poorly Poorly Fair Well Very Well

Why or Why not?

2. Did you think the elements and criteria included covered the key elements of your program?

Yes

No

Comments:

3. Was there anything missing you would have liked to be able to report?

Yes

No

If yes, please explain:

4. Were there elements that did not apply to your program at all?

Yes

No

If yes, please explain:

5. Please assess the Evaluation Instrument on the following:

	Very Poor	Poor	Fair	Good	Very Good
Instructions for use	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Number of criteria	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Rating scale provided	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Time to complete	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Comments:

6. How would you rate how easy or difficult it was to find information necessary to complete the Evaluation Instrument?

Very Difficult	Difficult	Neutral	Easy	Very Easy
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Comments:

7. Did you consult with or reach out to anyone else in order to get information necessary to complete the Evaluation Instrument? If so, whom?

8. Overall, how would you rate the ease of use of the Evaluation Instrument?

Very Difficult	Difficult	Neutral	Easy	Very Easy
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Comments:

9. Did you gain new ideas or insights into improving your program through the act of completing the Evaluation Instrument?

Yes

No

Comments:

10. Overall, how would you rate the usefulness of completing the Evaluation Instrument to your ability to evaluate and improve program quality?

Not at all Useful

Slightly Useful

Moderately Useful

Very Useful

Comments:

11. How do you think the Evaluation Instrument should be used in the future? Please choose all that apply.

- Teacher professional development
- Technical assistance to local programs
- New teacher training
- Program recognition
- Highlighting best practices
- As a program approval/review tool at the state level
- As a program approval/review tool at the local level
- To allocate funding

Other (please specify):

12. What changes, if any, should be made to the Evaluation Instrument to allow it to be used by more programs?

Program Information

13. School/Institution(s) involved in the Program of Study:

14. Name of Program of Study:

15. Career Cluster

- Agriculture, Food & Natural Resources
- Architecture & Construction
- Arts, A/V Technology, & Communications
- Business Management & Administration
- Education & Training
- Finance
- Government & Public Administration
- Health Science
- Hospitality & Tourism
- Human Services
- Information Technology
- Law, Public Safety, Corrections & Security
- Manufacturing
- Marketing
- Science, Technology, Engineering & Mathematics
- Transportation, Distribution & Logistics
- Other (please specify):

16. How long has the Program of Study been operating?

- 2 years or less
- 3-5 years
- More than 5 years

17. Name of individual submitting Evaluation Instrument:

18. E-mail:

19. Please provide your Perkins performance levels for each of the Core Indicators of Performance for this Program of Study.

Notes: You may need to reach out to a district CTE administrator or your state CTE office to request this data. Also, please ensure the data you provide is program level -- not the data reported for your entire institution or district. You may provide only secondary or only postsecondary data if you do not have both available, and please only provide non-trad data if your program has been identified as non-traditional. Please contact Alisha Hyslop at ahyslop@acteonline.org if you have any questions about how to report this data.

Secondary - 1S1:
Academic Attainment in Reading/Language Arts

Secondary - 1S2:
Academic Attainment in Math

Secondary - 2S1:
Technical Skill Attainment

Secondary - 3S1:
Secondary School Completion

Secondary - 4S1: Student Graduation Rate

Secondary - 5S1:
Secondary Placement

Secondary - 6S1:
Nontraditional Participation

Secondary - 6S2:
Nontraditional Completion

Postsecondary - 1P1:
Technical Skill Attainment

Postsecondary - 2P1:
Credential, Certificate or
Diploma

Postsecondary - 3P1:
Student Retention or
Transfer

Postsecondary - 4P1:
Student Placement

Postsecondary - 5P1:
Nontraditional
Participation

Postsecondary - 5P2:
Nontraditional Completion

Appendix F: Phase Two Interview Protocol

Call-in Information:

Call-in Number: 1-877-318-2398

Conference Code: 8708025792

Leader Pin: 3281

Name of Interviewee:		Date:	
Title:		Time:	
School/Institution:			
Program:			

Introduction:

- Summarize background on ACTE's HQ CTE project and personally introduce researcher
- Summarize the purpose of the research and results of Phase 1 of study
- Summarize purpose of this interview
- Provide assurances of confidentiality and consent to voluntary participation
- Secure permission to audio record the interview
- Start audio recording, repeat permissions/consent on recording

Questions:

**Actual questions will be developed after Phase 1 of the project and will build of information each interviewee provided in their Phase 1 questionnaire. Interview questions will be designed to dig deeper into the second and third research questions of the study, around the utility of the Framework evaluation tool and its implications for widespread use. A few sample questions are included below:

- How did you prepare to complete the evaluation (reviewing materials, consulting with colleagues, etc.)?
- What did you think about the length of time it took to complete the evaluation?
- Did you think that the evaluation instrument completion was a useful exercise? If so, what insights did you gain about your program?
- Are there additional tools and resources you would have liked to receive along with the evaluation instrument to make it more useful?
- Do you think that other CTE programs at your institution would find the tool helpful? Why or why not?

Appendix G: Results From Individual Element Correlations

Element		Technical Skill Attainment	Completion	Placement
Standards-aligned and Integrated Curriculum	<i>n</i>	20	23	20
	<i>r</i>	.53	.29	-.10
	<i>p</i>	.02*	.17	.69
Sequencing and Articulation	<i>n</i>	20	23	20
	<i>r</i>	.51	.47	-.06
	<i>p</i>	.02*	.02*	.793
Student Assessment	<i>n</i>	20	23	20
	<i>r</i>	.23	.50	.18
	<i>p</i>	.33	.02*	.46
Prepared and Effective Program Staff	<i>n</i>	20	23	20
	<i>r</i>	.09	.22	.19
	<i>p</i>	.69	.31	.42
Engaging Instruction	<i>n</i>	20	23	20
	<i>r</i>	.00	.12	.17
	<i>p</i>	.99	.58	.46
Access and Equity	<i>n</i>	20	23	20
	<i>r</i>	.56	.37	-.02
	<i>p</i>	.01*	.09*	.93
Facilities and Equipment	<i>n</i>	20	23	20
	<i>r</i>	.27	.52	.39
	<i>p</i>	.24	.01*	.09*
Business and Community Partnerships	<i>n</i>	20	23	20
	<i>r</i>	.61	.30	-.05
	<i>p</i>	.00*	.17	.84
Career Development	<i>n</i>	20	23	20
	<i>r</i>	.34	.38	.14
	<i>p</i>	.15	.07*	.57
Career and Technical Student Organizations	<i>n</i>	19	22	20
	<i>r</i>	-.32	.19	.03
	<i>p</i>	.18	.40	.90
Work-based Learning	<i>n</i>	20	23	20
	<i>r</i>	.31	.39	.25
	<i>p</i>	.18	.07*	.29
Data and Program Improvement	<i>n</i>	20	23	20
	<i>r</i>	.35	.38	.26
	<i>p</i>	.13	.07*	.27

* $p < .10$